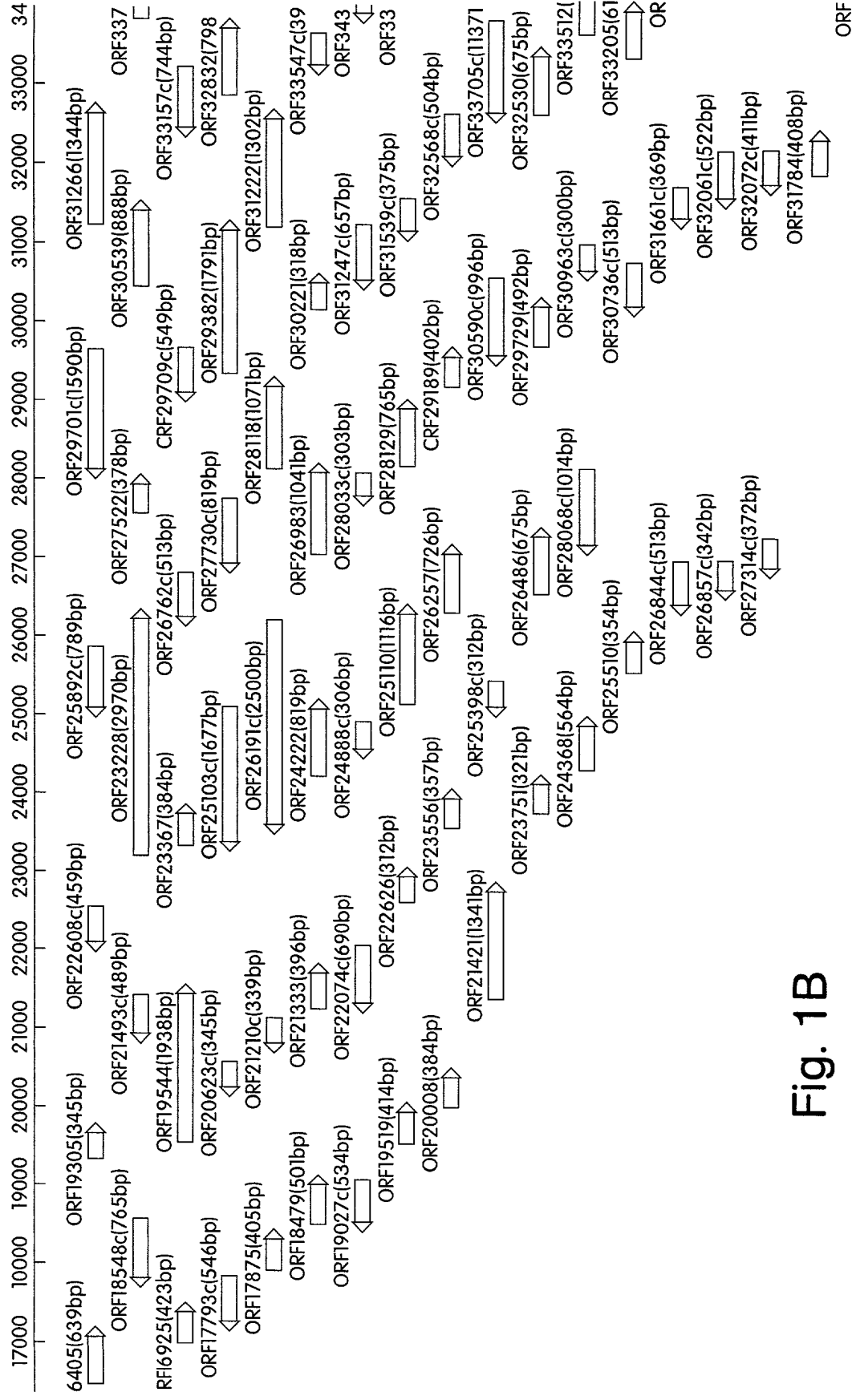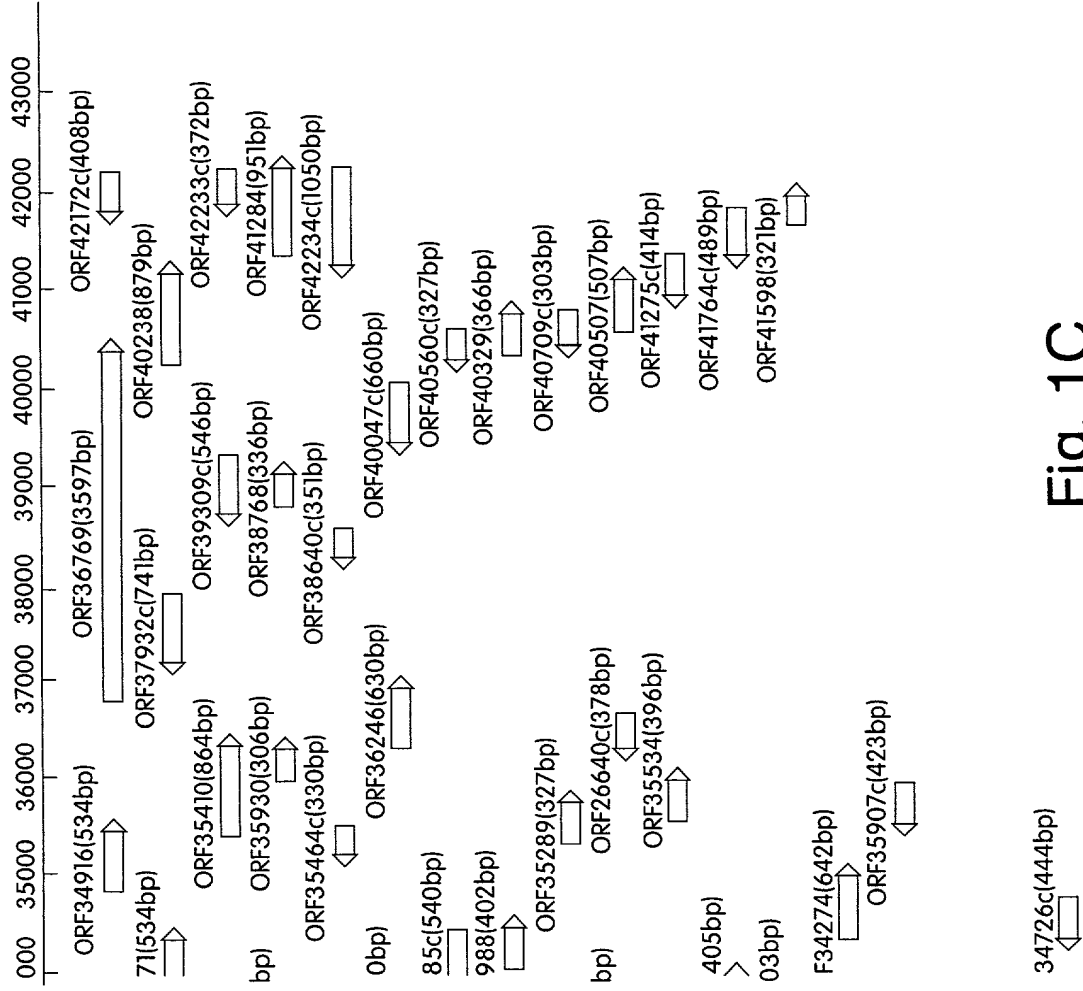Fig. 1A

Fig. 1B

Fig. 1C

BI48 SEQ ID NO:1

```
ATCGCCGATCCAATGCCAAGGAGTACCTGGGCAATCAGAGCCTACTCACGGCTGCCGGGGCCGGCATTGCCAAGCTCCTG
GACGCCGACGAGAACAACACCAGTACCGTCTTCAGCGGCAACGGCACCAGCTTCGGGACGACCGGAACCAACAGCAACTC
GGCCCTCAACAGCATCCTCTCCGGCGGCGTCAGCGACATCCGGCAGTGGATGAACAAGTTGTACGGGGAGGCCTTCGCCG
CCGTCTACGTGCAGCCAGGTGCGCGGGTCGCAGTGCATCTCGATCAGCAACTGGCGATCGACTATGAACTCAAGGGCCGC
AAGGTCGATTACAGCTCTGGAGCCGCTCATGCAACAGCAGACTTGGACTAACCCCCTTCTTCGTCTCTGCGCCGGCCTGG
CCTGCGCGCTGACCCTGGCAGCGTGCTCCACCAGCAAGGAGGAGATGCTGCCCCACGGCGAGGCCAACATGCTCGACGTC
TGGGAGCGAGGTGCGACCAGCTCGATAGGCAACAGCCGTGGCCGGCTGCTCCTCGATGCCAGGCAAACGCTGCGGCGCCC
AATCGATCCGCAGCAGGATGCCTCCGCGAATGACCAGGCCGACTACACCCGCACGGCCAGCAACGAGATCCACAGTCAGT
TCAAACGACTGCCCAATCCCGACCTGGTGATGTATGTGTTCCCGCACCTGGCCGGCAGCGATCCCGCCCCGGTACCGGGC
TACACCACCGTGTTCCCCTTCTACCAGCGAGTCCAGTACGCCATGCCGGGCGAACGCACGGAGGACTATTGATGGGCTTT
TTTCAAACCCTTCTGCGCGGTCGCACACAGCCTCAGTCGGTACCGGCAGACGCTCCCGAAGATTCAGGAGCGCTGGACGT
AGCGGCCGCGGAAGAAGCGACTGAGCGCTATCTGGCGCGACTGGCCGCCATGGGTATTCCTCTGCCCAACACCGGGAGCA
AGAATGGCGCCACGCAGGCCGAAGCGTCACGCCTCTACGATCACGACCCATCGTTCGTAGACCTGCTGCCCTGGGCTGAG
TACCTGCCCGACGAGCAAGTGATGCTCCTGGAGGATGGGCGTTCGCGCGCCGCATTCTTCGAACTGGTGCCCTTGGGCAC
CGAGGGCCGCGATCCCAATTGGATGCAGAACGCCCGGGACGCATTGAAAGAAGCCCTGCAGAACTCCTTCGACGAGCACG
AAACCTCACCCTGGATTGTCCAGTTCTACGCCCAGGACGAGATCAGCTGGGACAATTTCCAGGAGCAGTTGAGGCAGTAC
GTCCATCCTCGAGCGCGAGGATCGGCCTTCAGCGAGATGTACCTGGCGCTCATGAAGCATCACCTGGAGGGCATTTCGAA
GCCGGGCGGACTGTTCGTCGACACCGCCGTCAGCAAGCTGCCCTGGCGAGGACAACAGCGCCGCGTGCGGATGGTCGTCT
ACCGCCGGATCCGCAAGGAGGATGCGCAGATTCGCGGACAGGACCCGGCGGCGTACCTGAAATCCATCTGCGAGCGTATC
CAAGGCGGCCTGGCGAACGCCGGCATCGTCGCTTCGCGCATGGGCGGACAGGAGATCAGGAACTGGTTGATCCGCTGGTT
CAACCCGCACCCGGATCACCTCGGCCAGGCCGAGGCGGACCTACGTCGCTTCTACGAACTGGTATGCCGTCCGGACGAAC
CGATCCTGCAGGATGAATTGCCACTGGCCGACGGCACTGACTTCTCCCAGAACCTGTTCTATCGGCAGCCTGTTTCCGAT
GCCACCCAGGGCGTATGGCTCTTCGATGCCATGCCGCACCGAGTGATTGTGGTCGACCAGTTGAACAAAGCGCCGCTGAC
AGGCCACTTCACCGGCGAGACGCTCAAAGGCGATGGCCTCAACGCCCTGTTCGATCGAATGCCCGAGGACACGCTGCTGT
GCATCACCATGGTCGTGACGCCGCAGGACATGCTGGAAGGGCATCTGCAGCAGCTCTCGAAAAAGGCCGTTGGTGACACC
CAGGCCTCGATCCACACCCGCGAGGACGTGGCCACCGTTCGACGCCTGATCGGCCGGGAGCACAAGCTCTATCGCGGAGC
GATCGCTCTGTTCGTGCGCGGCCGCGACCATACCCAGTTGGAGGAACGCTGCATCACCCTGAGCAACGTACTGCTCGGCG
CCGGCCTGGTGCCGGTCGAACCGCAGAACGAAGTCGGACCGCTGAACAGCTACCTGCGCTGGCTCCCCTCAAACTTCGAT
CCAAACGAGAAGCGAGCCCTGGAGTGGTACACCCAGATGATGTTCGCTCAGCACATCGCCAACCTGTCGCCCATCTGGGG
GCGCACCACCGGTACCGGACACCCTGGCTTCACGCTGTTCAACCGTGGCGGCGCGCCGTTGACCTTCGACCCGTTCAACA
AGCTGGACCGGCAGATGAATGCCCACGGCTTCATCTTCGGGCCAACTGGCTCCGGCAAGTCGGCGTCCCTGACCAACCTC
ATCTGCCAGATGCTCGCCATGTACCTGCCGCGGATGTTCGTCGCGGAAGCGGGCAACAGCTTCGGCCTGCTGGCCGACTT
AGCCAAGCGGTTTGGCCTCTCGGTCCACCGGGTGCGCCTCGCCCCGGGCTCCGGCGTCAGCCTGGCGCCGTTCGCGGACG
CCATCAAGCTGGTCGAGAGCCCCGACCAAGTGAAGGTGCTGGACGCCGAAGACATCGAGGCCTCGGACTCGGTCCAGGGC
AGCAAGGCCGACCTCGAGGACGACCAGCGAGACATCCTGGGCGAGATGGAGATCGTCGCCCGCCTCATGATTACCGGTGG
CGAAGAGAAGGAAGATGCGCGCCTGACCCGTGCCGATCGCAGCGCCGTCCGCCAGGCGATCCTGGCGGCGGCCAGGACCT
GCGCCGCCGCGAACCGCACGGTACTGACCCAAGACGTGCGCGATGCGCTCTACGAGGCCTCCAGGAGCGATAGCACCGCG
CCAGAACGCCGCGCGCGGATCGCCGAAATGGCGGAAGCCATGCAGATGTTCTGCATGGGCGCCGACGGCGAGATGTTCAA
TCGCGAAGGCACGCCCTGGCCTGAGGCCGACCTTACCGTGGTGGATTTCGCAACGTACGCGCGCGAAGGCTACGCCGCCC
AGCTCGGGATCGCCTACATCTCGCTGCTGAACACCGTGAACAACATCGCCGAACGCGACCAGTTCAAGGGCCGGCCAATC
GTCAAGATCACCGATGAGGGGCACATCATCACCAAGCACCCGCTGCTGCTGCCCTACGCCATGAAGATCACCAAGATGTG
GCGGAAACTGGGCGCCTGGTTCTGGCTCGCCACCCAGAACATCGACGACATCCCAGCCTCCGGGGCGCCGATGCTGAACA
TGATCGAGTGGTGGTTGTGCCTGAACATGCCCCCCGACGAAGTAGAGAAGATATCCAGGTTCCGCGAGCTGTCGCCGGCG
CAGAAGTCGATGATGCTCTCGGCCCGCAAGGAAAGCGGCAAGTTCACCGAGGGCGTGCTCCTGGCCAAGGGCAAAGAATA
CCTCGTCCGTGTGGTTCCCCCGAGTCTCTACCTGGCCCTGGCCATGACCGAAAACGAAGAAAAGAACCAGCGCTACAACA
TCATGCAAGCCACCGGCTGCGACGAGCTCGAGGCGGCCTTGCAGGTCGCAGCGGATCTCGACAAGGCGCGCGGCCTGCCA
CCCTTCCCCATTGTTTTCCCAGACCAACCGGCAGTGGAGTGCCAGGACGAATGAGAGTTCTGAATTCGCTGACCCAGAAC
CTGATCGACAACCTGACCCAGATCCTGCAGAACCCCGAAGAGGATGCCCTGCAGACGCTAAGGATATGCGCTCCTGTACT
GATAGAGGAGCTGCAGCAGATTCAACTGAGGGCAGTCGATCGCCGGGTATATCGTCCCGCAGATAAAGCAGCTCTTGGATG
AATGGCTGCAACAACATCCACAGCCTGATACGGCCCAACAGGCGCTCATTGAGGCCGTGGACCGCGCGGAGATCCTACAG
CGGAGGCAAGCGTGAGACTCTTGAAGGGCGGCTGGGCAGCCAAACGATTTCAAGGTCCCGCCCTGCCCTGGGCGGGGCTG
```

Fig. 2A

```
CTGCTGGTCTTGCTGGCTGCATCCGCCGTAGGGGTAGAGCTTCTGGTGAAGGGCCTGCCAGCCAACCACAGCCTCTACGG
CGATGCGAAAGCGCGCTGGACGATCAATGAATACGCCGACCTGGAGTGCCCCTTCTGCAAGGTCTACACCCCGCGGCTTA
AGCGCTGGGTAGACAGCCATCCGGACGTGAACCTGGTTTGGCGCCATCTTCCCCTGCAGATGCATGGCGAGGCGGCCCGC
CACCAGGCTCGCCTGGTGGAGTGCGCGGGGATCCAAGGCGGCGCCAAAGCCTTCTGGAGCGCTATCGATGCGATCTTCGC
TCAGTCGGCCGGCAACGGGGGCGGGCTGCCTGGCGGCACATTGGACTTTCCTGAACTGGACCAGGCTCGACTGGAGAAAT
GTGCGAAAGACAACGAACTTATTGACTCAGATATCAAGTTGGACATCGACATTGCACGGTCGAAGGGCATTACAGCGACC
CCGACCCTCGTCATCCGGGACAACCAGACGGGACGAAGCGTGAAGCTTGAAGGCATGGCCGACGAGACCACGTTGCTGTC
GGCGATAGACTGGCTAGCCAAGGATCTCTAGCGTCGCGCCAAGAGACTCTTGGCTAAGAAATCGGCGAGGATTCCAACAT
CCCTCTTTTGGTCCTCCAGGATGCCCTGCACTTCACCTGGCAGAACCTCGACCTCCTCCCCATCCACAATCTTTACCATT
CTCTTGTGGCCGGAGCTGGTGAGGCTAAGCCTCAACTCCATTGCCGGCCGAGCATTGATGTAAATGCTCTCGAGCAAGCG
CTCCATGACTTCGACCACTCCTTAATATCAGTTAGCCAGCTACATACAGGAATTATGCTACCCAGGACATGCAGGCGTCA
CCCCTACTTATGTACGTGGCAGCGTTCGATCACGGCTCGAAAAAATACACCACCTACGAGTTGATGATTTCCTGCTGCCC
TATCGGAAGTGATCCTGTCTGCTGTCTGTACCTTTCTAGAACCGGTACAGACCCATGCCTCTTCATCACTCCCCCCCTGG
CCGGCGGCCACCAACGCTGGCCGTTGGCGTACTACTGGTACTGCTGAGCAGCGCGAGTCAGGCCGAAACCTGGGTCATCA
CCGACAAGGCTCATCCGGTCTCTGCCACCGGATCGTCGCGCGTTCTGTTTCTGGACGCCCAGGAACACCTCGAGGAGCAA
CTGACTGCGGCCTTGCCCCAGGATCCACAGCATGCTCAAGCGGCGTTTAAGCGATTGCTACAAAGCCCCGATGGGCGCCG
CCTGCAGGCAGAGCTGGTCAAGGCACAACAAGACGTCGCCGATGCGTGGAGTCTCGGTGTCGAGAAGATCCCTGCCGTAG
TAGTCGATAGGCAGTACGTGGTCTACGGCGAACCGGATGTTTCGCGCGCTCTTGAGCTAATCGCCAAGGCCAGGAGGTCG
CGCTGATGACCAGCCTCAACCTCCGCCGCCTGGCAGCGGCGGCCGCCACCTTCAGCCTCTCGTTCACGGCCTCGGCCGCG
ATCAACAGCGCTGCCATCGTCTCCTCCACCCTTTCCCCTCAGTGCCTCGAATACAAGGTCGTCGGGATCTGTTACTGGCT
GCTCTGCGGCCCGCATGGCTGCAAAGTGAAGACGTCGGTCAAGGTCCGCCACTACGTGCCTGACGCAGTCGTCTCCAGCT
ACGCGAATACCGGGAGCAACCCCTGGACCGAGGTATCGGCGCTGGGTACACCGAATCCACTCGCCCAGGCCGGCAATGAC
GCGACCACAAACTACAAGGCCGAGAACAGCATCGGCCGCTTCAAGGAAGCGGATGTGATCGGCCATCCTGGTGGCGCCAC
GTTCAGCCGGTTCGCCAGCGCCTCTGGGTACGTTTGCCCTGGCGCCACCGTCCCGCTGGTGCCGTACTTTCTCAGCACAC
TGGACGCCATTGGCTGGCGGCATGGAATTCCCGAGCAGGTGTACCCCGAAGCGTTGGTCCCAGGGCTGCGCGAGGTGGGT
GGAATCTTCTCCGGCGACATGTGGGGGAACCTCTATCCGCGCAGCGGCTTCCTGCACCAGACCGACGACTACAAGACGGC
AGCCGTCATCGCCCAGCGCGCCGGCGATATCACCACGCGAATCGGCCAGCTCCACGTCTACCTCCCCATGCGCGCAGCCC
CCAAGGACGGCTACTGGCCGGCGGGCGAGCTGAAAGAGGGCGATGCCTCGACCGGGAAATGGCAGGAGCTGACCCCATCC
CTGAGCCTCAACTGCGCGGTGTTTCCCAACTCTGGGCCGAAGACGCAAGCCGTCGACGGGGAGCACGCCTGGGCGCTCTG
GCGTCCCTACTCCTGCTGCCAGCGCAAGGGGCAGATGTTCATCTGCAGTACCGACTTCCAATAAGGACACGGAGACGAAT
CATGCGAATGAACATCACCTCGGTCGCGCTAATGTGGCTGCTCGCAGCGCAACTTGCCCAGGCCGACGACCCGATCAACG
TGTCCAAGACCGGCACGGTGCTCAGCGACGAGGTCCTCTACAGCATTGGCGGCGGCAGTGCGGTGAGCATGGGCAGCGCC
GGCCAGATGGACTCGATCGGCGTCGGCTTCGGCTGGAACAACGACATGATGTGCGGAAACATGAACCTGAGCACCACCCT
GGAGAACCAGCTCAACGGTGCCACACAGGGTTTCCAGAACATCATGGGCTCAGTCATCCAGAACGCGACCGGCGCGGTCA
TGTCGCTGCCGGCGTTGATCATCCAGCGCGCGAACCCTCAGCTCTACAACCTGATCACCAATGGCATCCTGCAGGCGCGG
ATCGACTACGACCGCTCGAAAGGGACTTGCAAAACGATCGCCGAAAAGATGGCTGACATCGCTGGCGAGCAGACCGGCTG
GGGGAAAATCGCCGAAGGCCAAGCCCTGGGCGCCACACTGGCCTCTGACGGGAAAGACGCCGTATCCGCCCTCGAAGCAG
TGGAGAAGAAAGGCGGCAACGATGGCGTAACCTGGGTTGGTGGAGACAAGGCCGGCGGCTCCGGCCAGAAGCCCATTCGC
ATCGTCAACGACGTGACCCGGGCGGGCTACAACCTGTTGACCAGCCGCTCGGTGAATGATTCGTCGAGCGTGCCTTCCGC
CACTTGCAACAACGGCCTGGTCTGCAACACTTGGTCCTCCCCCCAGGAGGCCGCCGCATTCGCCACCCGGGTACTGGGGG
AGCAACAGCAACAGACCTGCGAAGGCTGCCAGAAGACGGTGACGGCTGCTGGCGTCGGCCTCACCCCGCTGATCCAGGAG
ACCTACGACAAGAAGCTCCAGTCGCTGCAGGAGCTGCTGTCGAAGAGCAAACCACTGACTGCAGAGAACCTGGCTGCGGC
CGGCACCGATGCTCTGCCAATTACCCGCGGCGTCATCGAGGCGCTGCGCGACGAGCGTGACCAGGACGTCCTGGCGCGCC
GCCTGGCGTCCGATGTCTCCCTGATGGACGTGCTCAGCAAGGCACTGCTACTGCAGCGCCTGATGTTCGCCGGCGCCAAG
GAGCCCAACGTCGCCGCCAACGGCCTGGCCACCCAAGCCGTCGATCAGCAGACCAGCCTCCTGCAGCAGGAGATCTCCAA
TCTCAAGACCGAACTGGAACTCCGTCGCGAGTTGGCCAGCAACTCCCCCATGCGGGTCATCGAGCGCGGGCAACAACGCG
CCTCAGGGTCCAGTGGCGTGTTCGAGTCGGCGCCCGATGCCGATCGCCTCGATCGCCTGCAGGCCCCCTCTGCCGCCGGC
GGCAAGTCGGGAGGGAGACCGTGATGGCAGATACGCTCACCACCCGAAAGCTTCTCGGTCAGCTACTGGTCGGAGTGCTG
ATCGTCATCGGACTGGCAGTGGTCGGTACGCTGCTCAGTCTCTTCGCCCTGAACCACTTCGGTGGCATCCAGGGCCTGGA
GGCCTGGCGGCAAAGCAACTACTGGAGCTTGTTCGCCTTGCGGGGCGCTGCTGTACTGCGCCCTGGCCATCGCCTGGTTCC
GGCAGCGCAAGGAACTGAGCGCGCATGAGCGGCAGCGCATTCGGCGGATCGAGATCCTGGTGCTGTTGCTGGTCCTGCTC
```

Fig. 2B

```
ATCGAATTCAGCAAAGCCTACTTCCGCACGGGAGGCGCAGCATGACCTTCATGACCAATGACTACCTGGAGTATTACCTC
ACCCTCCTCGGCTGGATCATCAACAACGGGATCTGGAACATGATCTCGGATACTGGCCTGTTCGCGGTGCCGTTCGCGGC
CATCGTGATGCGCGAATGGCTGAAAGTTCGTGGGGAAGGCGCCGACGAGGGCAACAAGGGAGTGCTGTCTCTCGCCCGCA
TCGAGACGCATATCTACGTCGGCTACATCGTGGTCGCCCTGGCGGGGATCCCGGTCGTCAACGTGAGCTTCGACACCATC
GAGTTCGACCAGACTCGCGCCCAGCAGTGCCAATACAATCTGCCGGCACCGGCGGACACCGGCTGGTCGAGCTCCTTCAG
CAGCCTGGCCGGCAAGAGTGCGCAGATGCCGCTCTGGTGGGCGATGATGCACGCCCTGTCCAAGGGCTTCACCAGCGGCG
CCATCGCGGCCATTCCGTGCGGCACGGATCTGCGGCAGATGCGAATGGAAGTGGACAACACGCGCGTGAACAATCCGCTG
CTGGCACAAGAAATCGCTGATTTTTCCAGAGACTGCTACGGGCCTTCCCGTGCGCGGCTGTTCATGCGGCAACCCGACCT
GGGCTCCGTCGCCGAGGACAACAAGGCGTTGCAAGACCTGAACTGGATCGGCTCCCGATTCTTGTTGAACACCCCGGGGT
ACTACGACACCGACTACTCGAAGAGTCCCCGTCAGTCGTGGCCCTACAACGCCACCCGCGATGCCGGCCTGCCTCAGGTG
GGCGGTGGTGGCGGCTACCCAACCTGCAAGCAGTGGTGGGCTGACTCAGGGATCGGCTTGCGTGATCGGATCAAGGACCA
GGTGGATCCGGACCTGATGACCAGCTTCCTCAAGTGGGCGAAATGGTTGAACCAGGACGAGGTGACCGAGGCTGTCATTC
GCCAGGTGATCTCACCCTCCAGCCAGGTCAAGGGTAACGTCTACACCGATTACGGCGGGCAGGTGGGCGGCACCGTGTGG
AACGGCATCGCGAGAACCGCAGGAACCTTCGGCGTTGCGGTGGGCAGCTTGGCATACTTCCCGGCGATGGATATGGTCCG
CCAGGCACTGCCGATGGTGATGTCGTTCCTGAAGATGGCAATGGTCATCTGCATTCCGATGGTCCTGGTCATCGGCACCT
ATCAACTGAAAGTTGCCATGACGATGACGGTCGTCTTCTTTGCGATGATGTTCGTCGACTTCTGGTTTCAGTTAGCCAGA
TATATCGACAGCACGATACTTGATGCTTTCTATGGTTCGGGATCACCACATCTTTCATTCAACCCAGTCATGGGGCTGAA
TACGGCTACTCAAGATGCGATCTTGAACTTCGTTATGGGTTCTATGTTCATTGTTTTACCACTACTGTGGATGACAGCGA
TCGGCTGGTCCGGAATTCAAGCAGGGTCTGTTCTGAACGGATTGAGCAGAGGGACTGAAGGAGTTCAAGCCGCCGGCAAG
GAAGCAGGAAATAGAGTTAAAAACGCAGTTTGAGGAAGTATAAAGCCATTACCGGGCTCTAGTCCCGGTAATGGTGTGCT
GCCTGATAGTCAAACAGCTCTGCTACAGATTCAGTTTCAGCTTGTGAAGTCATAGTTGAGACATCTCGACTTGGAGACCC
TTTAGAGACGGAGATGTCCCCTTTAACCAAGGCAAATCCCACCAGAAAGACACACAGTAGAGTCATTGATACCCCAGCAA
GCCACATAAAATTCTGTGACACTACAAGAGCTACTAACAAGAATAGGCTAACTCGCTTAACCCAACGTAGAGCTGGACGG
CGATCATGCAGGCAGAAACGCACAAGCATACCCAGACCAAAACCGATCCGGGAGGCAAAGCCTTTGTTGGTGTGCGCGTT
CATCATCAATCTCCTGGCTCCCAAAGGGAGGCATCCTGCTATCACCTATACGCCGAAAAAGATGATTTGGCAAGCATTAT
GGCATATTATGCCACTAGCTATCTGCCGACTGGAGTACCTCATGGCAACGCGAAACGTCGTCCTTCCCGATCCGCTGGAG
CAGGATATCAACGAGCTGGTGGAGACCGGCCGCTATCAGAATCGCAGCGAAGTCATCCGGGCAGGCTTGCGCCTGCTGCT
GCAACAGGAAGCCCAGATANGCGCCAAGCTCGAAACCCTCCGCAACGCAACATCCAGTGGGCTGATGCAACTGGAGCGCG
GCGAGTACGACGAGATCACCAGCGACGAACTGGCCCAATACCTCGACGAGCTCGGCAACCAGGCGAGCCACTGAAGCATG
GCCAAGTACCGCATCTCTCATGATGCCCAAGCGGACATCGTCGATATCCTGCGCTTCACCCACAACCACTTCGGCGATGC
CGCGCGCCGACGTTACCAGGCACTCATAGGGGCGGCGCTGGAAGCAGTTGCGACAGACCCACAACAGGTAGGCAGCATCA
GCCGTGAAGAACTGGGAGCTGGCCTGCGCAGCATCCACCTCGTTTACTGCCACTCGATGCCCAATGTCGGTAAGGTTGTT
CGGCCCAGGCACTTCGTCTTCTACCGGGTGGCGACAGACCAGGTGCTAGAGGTGGTTCGCGTGCTTCACGACGCCATGGA
TGTGGATCAACACCTGCCCCAACGATGAGCAACTGTAACAGGAGCAAATGGCCCAAAGGGGAAATGGGCTTGAAGGGCGA
AGGGAATGGTTCACAAGGGTAAAAGCCCTACCCGAAAGGGCCTGCCGGAAGGCAGAAAGGGCTGCGTTCGCGCCGGCGAA
CACGAAAAGGCTACGGGTGAACATCCACCAGCAGGTGCCCCATGGTTGATGCGCAATAGCAGTGGCTTGACCACTCGATT
GCATTCCCCCTGACCAAGCATTTGCATTCGTGCTGACCACCGCTTGCATGCGGCTTTACCAGTCACACCGTAGTCACGAC
CGGCTGCAGTCGACCCAAAGCGGACCTTTGTGACAGACCGAAATCGGCCAAAAGCAGACGCTCAAAAGCATCTACGAAAG
CAGGGACAATTCCGGGGGATACTCCTCTACGTACGTTGCTTATATCGCGGAGGAATTGGCTGAACCGCGGTGCGGAGAGA
TCTCCTCAAACTGATGGGTTGCACGCATATCGAAGCAGATTACATAGGAGGCTTGCGCTGTTCAACAGCTCCTGAGGGGA
CTTGGGTTGCCCATGGTTTCCACGGCCCAATCGTTGACGTCATTGACGATTCCGCTGGCTTTTTTCAGTACGCATCGCTTG
GCGCTCCATTACCCAGCCCAATGCGGCCTTGCCGTTGACCAAGCGATTCCAAGGACTGCGATCCATGTAGCCAGCCCTCT
AATGCATGTATGTATAGGTAAGGTCGTCGTTATTTCGGCGTGGATGTGCTGACTGGAGGTTTAGCGTGGTGAGTCAGTAT
GTCGTGCGCTATATAAGTCTCATCTCGGCATCCGCACATGGATGCTGTTAGATTCAGCAAGTGGGAAAAAGCGCGAGCGC
TAAAAGGAAAGCTGCCGGCCGGCGTGCTCAGTAGCAAGCTACCCCGGCAACACTTAGGAAGAGGTGATCATGAAGTTACA
GGCATATCGGCTGCAGAACTACCGCCGGCTGCGCGATGTTGTCATCGAGCTCGATGACGAAATTTCTATCTTTGTCGGTG
CCAACAACAGCGGGAAGACATCCGCCGTCCAAGGCCTGTACTCAATGCTTCGCGGCGAAGTGAAGAAGTTCGAGCTCTTT
GACTTCAGTGCGGCGCTGTGGGCCGAGATCGATGCGGTCGGCAGGACGCCCCCTGGCGATGAGGATGCGCCCAAAAGGTT
ACCGTCCATACTCTTGGATCTCTGGTTCCGCGTCGGTGAAGACGACCTCGCCACTGCGATGTCGCTGCTGCCGAGCACTG
AGTGGGACGGCAAGTGCGTCGGGATCCGGGTAGCGTTCGAGCCTCGGGATGCCCACGAGCTCGTCTGGAAGTTCCATGAA
CTACATGAGAAGGCCAACAACGCAGCTGTCGCGCTTGCGGCCAAGCGCAAGGCCGCCGGGGAGCAAGCTGTGGAGGCGGG
```

Fig. 2C

CGCGGAAGACGCGGCTGCGGTGGTGGCCGATGCCGGCGAGTACAAGCCTTGGCCAGAAAGCCTGACGAAGTACCTCACAA
AGGAACTGAGCAAGGAATACACCTTCCGCTACTACGTGCTCGATGAGCGGGCTTTTGTCGGCTATCAGGCAAGGGAGGCC
GACTACGAGCCGCTACCCCTAGGCAAGGAGCCGGGCGGTGCAGCCATTCTCAAGTCGCTGGTGAGGGTCGACTTCCTGCG
CGCGCAGCGGCACCTCGATGACCCAGATGCCGGTAGCTCTGATCGCGCAGAGAGCTTGTCGCGGCGTCTGAGCAGGTTCT
ATCACCGCAACCTGGAGAAGCGTGGCGACGACCATGCGGCTCTCAAGGCGCTAGATACCTCGGAGAAGGAGCTGAACTTC
CACCTGAAGGAAGTCTTCAATGACACCCTCACGCGCCTGGCCAAGCTCGGCTATCCGGGCGTCAACAATCCGGAGATCGT
GATTCGGGCGGCCTTGGATCCGACCACTGTCTTGGGGCAAGACGCCAAGGTTCACTACGTGATCCCGGGCGTAGCTTCCG
CCCAACTGCCAGACAGCTACAATGGCCTGGGGTTCAAGAATCTGGTCTACATGGTGGTTGAGCTGCTCGACTTGCACGAG
CAGTGGAAAGCCGAGGATGACAAGCGAGCTCCGCTTCATTTGGTCTTCATTGAGGAGCCTGAGGCGCATCTGCACGCGCA
GATCCAGCAGGTCTTCATCAGGAACGTTTTGCGCCTCCTTGAGGATGCTAACGATCACGCGACTTTGTTCCACACGCAGC
TCGTCATCACCACGCACTCCCCGCACATCCTCTATGAACGCGGATTCTCGCCCATTCGGTACTTCCGCCGCGTCAACGAC
CAGTTGGGCCATCACACGGATGTGCGCAATCTGTCGCTATTCAAAACGGGCGCGTCCGACGCTCCAGCGCGCGAATTCCT
GCAGCGGTATCTGAAGCTGACGCACTGCGATCTCTTTTTTTCCGACGCGGTGATATTGGTGGAAGGCAACGTCGAGCGTC
TGCTCCTGCCTGCAATGATCGAGTTGGTGGCCAAGCGCCTGCGTTCTTCCGCCCTAACCATCCTTGAAGTCGGTGGTGCG
TTCGCGCATCGGTTCCAGGAGCTGATCGCCTTCGTTGGGCTCACAACACTGGTCATCACGGATCTGGACAGCGTGACGGT
CAAGACGGACGCCGAGAAGGCCGCCGCGCAAGGCGCAGGCGCTGAGGGCGCCGTTGACGGAGATGACGAGGACGAGGACG
ACGACCTGAAGCCCTTCGAGCTTGAAGACGACGACGAAGCAGAACCGAGTGGCAAGAAGAAGTCCAAGAAGCGTGGCAGC
ACCTGCCATGCACACGTGGAAGGTGCCGTCACGTCCAACCAAACCCTCATCAGCTGGATCCCGAAGAAGCGGTCGATGGC
AGAGCTCTGGGAAGTCACGGCGGAGCAAAAGACGCTGTCGCTGGCTGAGGATTCCAGCGCTGGGGTTCGGGTAGCTTACC
AGACCAAGGTTTCGGTGACGGTGGGTGCGACGACATCACAGCTCTGCGGCCGCACACTTGAGGAGGCCTTTGGTCTTGAG
AACGCGGACTGGTGCCAGGCTGAGGCAAACCGGTCGGTCGGCCTCAAGCTCAAGCGCGCACCGAGCAGCCCTGAAGAGCT
GGCTGAGAAGTTACACGATAGGGTGGTCGGCAAGAACTTCGACAAGACCCGCTTTGCGCTGGAGGTACTCGCAAGCGGGC
CGCTCAATGGCTGGAAGGTTCCCGCGTACATCGCCGAGGGCTTGGCCTGGCTCGAAGCCAAAGTGGCCCACGAGCTTGAG
GCGGATGCTGCCATCGCCACCGAGGTCGCGACTATTGAGCCGACTACAGCCGATGTTGTCGCTATCATTGTTGACCCGGG
GCAGACGGCATGAGCAGACGAATTGATAGCCCAGATACCGACGCCGACCGCGAGATCCACGCATGCATTGTAGCGACGCC
TCCGCAGCCCTTCGTGGTTCGTGCTGGCGCAGGTTCCGGCAAGACCACCTCCCTCATCAAGGCGCTGGACTGGGTGATCT
CGGAGCACGGCGCCAGCATGCGGGCGAGGAAGCAGATAGTCGCGTGCATCACGTATACCGACCTTGCCACCAATGAAATC
CTGGCGGACGTCAACGATGACCCGCTGGTTCATGTCTCGACCATCCACAGCTTTTTACTGGTCTATTGCAAAGACGTTCCA
GGCCGACATCAAGGTTTGGCTGCAGAACGACATCCGCAGGCGGATCTCCGAACTTGAAGAAGAGTTCGAGAATTACAGCT
CGCGTGTCCGGCAGACCACGCGCGACAGGAACAAGGCCGACCAAGAGCGATATGTCCGAAGCCTGGAGGCTGTGGCCGGC
GTCAGGACGTTCAACTACGGCGTGGGCAGTGACTACGCCAAGGGCATACTTGGCCACGAGGACATCCTTCAGCTCGCCGA
CTTCCTGCTACAAAACCGCCCGCTGTTCCGACGGGTCGTGGCGCTGAGCTACCCGTTCGTGTTTATCGATGAGAGTCAGG
ACACGTTCCCGGGTGTAGTGAAGTCTTTCAAGGAAGTGGAAGCCCAGATGCAGGGCAAGTTCTGCCTTGGTTTTTTTCGGC
GACCCGATGCAGTCGATCTTCATGAGAGGCGCAGGGGACATCCAGCTTGAGGATCATTGGCGGGCCATCACGAAGCCGGA
GAACTTTCGCTGCGCCAAGCAGATCCTTGACGTCGCCAATGCCGTGCGCGCGCAGGGCGATGGCATGGAGCAAGTCCGCG
GGCTGCACGAGAGGGTCGATGGGAACCTCAAGCTGGTGGAGGGGTCGGCCCGGATGTTCGTCTTGCCGAACACGCTGAAC
CGAACCGAGGCTTTGGCAAGAGTCCGAGCGTGGAGCTCGGCGACGAACAACGACGAGGGTTGGACAACCCCAGACATCGC
AGTCAAGATTCTTGTCATCGTGCACCGCATGGCCGCAAACCGGCTTGGCTTCGGCGGCATCTACTCGGCGCTGAACGACA
AGACGTCGGATGCCATGAAGCAAGGGATGCAGGACGGCACCGGTTGGCCCGTTCGACCCTTCCTAAGTTTTGCGCTACCG
ATCGTTGCAGCTGTGAAGGCCGGCAATGAGTTCGCGGCGATGAGCCTGCTCCGGGAATTCAGCCCGCGCCTGGCGCCTGC
GGCTCTGACCGGCCGACGTGCCGCGGATGTATTGCGAGAGCTGCACGCTGCTGCGTCGAGGCTTGTCGCCATGCTGGACG
AGGCAGGGACCACCATTGGTGACATAGCTCTCCATCTCTGTGACACGGGTCTTTTTGAGTTCGACGAGCGCTATGCGCGT
GTTCTTGGGTTTGTCAGGGATATTGCTGACACCGCTCAGGAGCCCGAGGCTGCTGATGCAGTTCCGGCCGAAGGATTATC
CTTGGACGCGACAATGGCCAAGTTCTTCAATTGCTCTGCGCAAGAGCTTTGGCCCTATGAACGCTATGTCTCAGAAGGCT
CCCCCTATGCCACGCAGCACGGCGTGAAGGGAGCGCAGTTCGAACGCGTCATGGTGGTGATGGACGAGGAAGAAAGCGAC
TACCGAACGTACAACTACGAGCGTGTCTTCGCGAGTGCTGAGGCCCGCGCTGCAGATCGTGCACGAGCACTAGACGGTGA
TGAAAACACTTGGAGCCGAACGCTGCGACTGCTTTACGTCTGCTGCACTCGTGCCCAGCGGGGGCTGGTACTAGCGTTCT
TTGTCGCCGACCCTGCGACCACCCTGGAAAACGTCGTGGCGAGCGGGATCTTGCCGCGAAGCGCAGTCTTTACGCAGGAA
GTGTTAGTTGGATGGCCATAGACGAACGATTGCCTGACCTATTTGAAGCCAGAGCCGAGCGGCAGGCTCGGTAGCTAGCG
GGTGGTAACTCTGAGGACGTAATCGACTGCAATTGGCCGATTTCTGCCTATGACGAGGGGCAGCTACCGACCCATTGCTA

Fig. 2D

TCCTTCGTAAGGGTGGCTACGCACCGGGCACGCTGCGGGATCGGCTGAAGCTAAAAAATCGCACAACAGCTTTGTAAAAC
AAGATCTCGGTCGATAACATCTATGAACGGTCTGGGTTGGCAATCAGTATTGGCTGCCAGTTCTTGTGACGATGCATGGT
TTAAATCGACTCTTTGAGGAAATGCGTGGGAAGCCTGGAAAAGTGCTGTTTCGCCTGCAAAGAAATAATTCATGTTCATG
CGATTCGTTGTCGGCAGTGCGGCGAGTCCCAAGGCTGGCGAAGGTTCATGAGCTCTCCAACCTCAGTAGTTGCGTTGGTC
CTTAGCCTTTTATCAATCGCTGCCACAAAACCTGTGGAGCGATTGTTCGATGCCCAGCGAGCAGAGCTACAAATCTCCAT
CACGGGTGGTGATTACAAAGCTGCCCAGCTTATGTTGACCAATAACGGGTCAAAGCCTGCAACTTTAGTTTCCTTCGAAA
TCACATCGAAAGCCACGACCAATACGAAAACATGGTTTTTGGTAAGCAATACGGATGGCGAAATTCTGGAGCCAGGCAAA
ACTTACAAAATCAGGGCCTCAACCGATGAGTCTATCCCAAAAATTGTCGAAGCTGAGCGTCGGACGATTTTGAAGTCTCA
GTACGCACTTGCAGATAATTGCGAATTAACCGCTAAATACATAGAGGCCACGGGGCAGAAGGTTGTGCGTGTGCAACCGT
TCATGTGCGACACACCTCCTGAAAAGGGTGGCCTGCCCCCTGGTAAACCTGGCATACCCATTTGGTACCTTGGTCAAGAA
TGATGTTTTTATGCCGCCCTGGGCTTTGACGCCGATTAAGCAAAGCTGTGTTCGCTCATCCAATACGTCCCTCGCCCAGT
TAAACGACTGTTATGTATATGGGTGCTGCCGCTACGTAATACCTTGGCCCTACGCATACGAAGTTAATTCTGAAAGCGTT
CAATGGACAATCTTCCTCCTCGGCGTCGACTGCAGCGGTAAGGTGATCTACTTTCGAAACACTGCAAGGGTAGGTCCTTT
TTTGGCAGCGTCCATATACCGACCGTGGTATGGCTCAGATGCGCTGGTACTGCATTTCACCAAATAAAGTTGTGCTATAT
CGCTCACGGCCGGTATTTCTATCGTCCAAGGCGACATATTGACGATTTCAAGGGGTTACGTACCCGTCGAACCCCGGCTCT
ACCTGACCGTCATTACAGGTTACCTGCACGAACAATCGCGCAAGGGTTCCTTTGGTTTGCACCAGACCAAAATAATCTAA
CGGTATTTTGTATCGATGTTTACTGCAGGTGATGACTTGCTCACCGGATTTCAGCACCAAATTTTGCTGAATTTCTTTTTG
GTTCCAAAAAAACGTCGACGGGGCTGGATGCGACCCATATACAATGCTTGCAGAGGCGGGCAAAGATTTTGGCTCGTAA
TATGTATGCCCCAATTGAATCTTGAGAGAAAACTCGTCAAATAAAGTTTTTTCACATATAGCGCATTGCGCAACAAGATG
TTCTAGATTTTTGTCAATGACAATCATTTTGGCTTAACTTCGTTCGGCGATGGCTGAGCGTTTTGTTGTGACTGAGGAGG
CTGCGACTGGTGTTCCAGCACAACCTTGTAATAGACGGCGTACATTATAAGAGAAGACACCAGGGGAACTTCCGATTAAAT
TCAGCGTCCACCCGACCCATATGGTGGGATTGAGCGAAAACCACTTCTTCAATCGCGCGCGCCGCGTGATGGGAAGCGTC
GATATTTCACTGCCACGCTTGATGGGGCTTTGGGTGTGCGTCGATTCTGTGGAGAAAAAAATCACAGAGCCAAGCTTCTC
GCCTTTGGTAAGGCAGACATTTTTGTTGGTGGTGTTGAATATCCTGAGCTTGAGCTTGCCATCAAATGCAGGTTCGACCT
TCGCCGAAGCAACCAGCACGCCGCGTGAGAGAAAAAGACTTCCCGTAGGTAGGACTATGCCGTACCGATTGTGCGGCACC
TTGATTTCTTCAGCGGCCTCCACCACAACAGAATCGTGACCCCTGATCGTCATACCATCGACAATGCGCCACAGGTTTTT
GTCGTTGCCAGAATAGTTATCACTCCACCCTTCACCCACGGACAATTCGAGGGAAAAACTCTTCAATTCTCTCGTCCCCTG
CATCAGTGAGTATAAGGCTGTTGCTGGAGTACGATGCCGCGTCAAAATCCGTGTGGGATTTCGTTGTACGCCCTTTGATC
TGTAACACACTCATCAGATTTGTCCTTCGTCCTTGGGCTGCGAGACGCGGCCACCCTGCCATTGTCCTTTATACCGGCCG
ATATCCCCGGATACCGCCTGAAAGATGACGTGCGCAAAGCGTGCACCAATCTGAATTTCAAACGCCTCGCTGTGATTGTT
GGTGAGCGCGAACGTCATCGGCCCTACATAACCTGGAGGCAGCACACTGGAACTGAACGTTATCCCGCTTCTGAACAGCG
TGCTTCTCGGAAAAAAAAGCGCCGCCAGATCTTCCGGCAGATCGAATTCTTCCATGGTGCTCGCCAGATAAGTCTTGCCC
GGTTCCATGACGAAGCAGTCATCCGGGTCTGCGAGCACGACCTCGCTGGCAGGGGTGCGTCGCGTAGATTCTCGCAAGCT
TCCACCCCCTACTGTCAGGCGAGAGAGGCCTGCGAGTCTGAGGTCAAATCCAACGCCTTCCGGGGTGGTCAACTCACGGT
GGGCAAGGTGCTTGATTAGTTTTCCATCCCGGACCAGTTTCAGGAGCGAGTGCGGTGAGTAAATCATCTATTTTGCCTCG
GGAAGGGCTCAGTCTATAGCACTGAGCAACCGCCTACGTTTACAGGCAAAAAGTCGGACTCAGTCGTTGTGGCCGCATTT
TGCGTGTCCTGCGGGTTTAGTAGCAAGGACAAAACCGTGCGACATGCGCAGCGCACTTTCACTGCACGGACAGCCCCATC
GGCATAGCAAGCCTTCACCGCCACTCTGATTGGGCTCACGTCCGACAAGAGTCGAGCAGCTTTGCAGGATTTTCCACATC
ACCAAAACTGCCAATGGCAGCTAATGGCCGTTCTCTGCCTGTCGCCTCTTTGGCATGACTGGTCAAGTCGGATGCAAACG
GTGGTCAGCACCAATGCAATTGGGTGGTCATGTGCGATGCAATTACGCAGTTGAGCCTGGCCCAGTTCCTCCCAAGCAAA
GCATAAGACCAAGATGGCACATTGCCAACAAAATACCCTTCCCCGCTACCGTTGTTTTATCGTTGTTGCCAGCCCTGATC
TGGCGGAAAAGCCCGCTCCATGAATCGTCATGGAGCCTCCCATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCCGTCC
CCCACCCCAACAACCAAAGCTGCCCCAGGGGGATTCATCCTTCCTCTGAGCAGCATGGAACTGCTCGGCACGCCTCGCCG
CCGGCAGCTACTGGAGAACATCTGGCAGCGCGCCTCGCTATCCAAGCAGCAATTCGAGGAGATCTACCGGCGGCCACTGG
CCAACTATGCCGAGCTGGTCCAGCAGCTCCCTGCTTCGGAAAATCATCACCATGCCCATCCAGGCGGGATGATCGATCAC
GGCCTGGAGATCGTGGCCTACGCACTCAAGGTACGGCAGACCTACCTGCTCCCGATCGGCGCAGCGCCGGAGTCACAGTC
AGCCCAGGCTGAAGCCTGGTCGGCCGCCGCGGCGTATGGCGCCCTGGCTCATGACATAGGCAAGATCGTCGTCGACCTGC
AGGTTGAGCTACAGGACGGCAGCACCTGGCACCCTTGGAACGGACCGATCAACCAGCCATACCGCTTCAAGTACGTGAAG
TCCCGCGAATACCAGCTCCACGGCGCTGCCTCAGCACTTCTCATCCACCAACTGCTACCGCGCACTGCACTCGATTGGCT
CAGTCGCTTTCCAGAGCTGTGGGCTCAATTGATCTACCTGTTCGCTGGGCAGTACGAGCACGCCGGGATCCTCGGCGAGA

```
TCATCGTGAAGGCAGACCAGGCCTCAGTTGCACAGGAGCTAGGAGGCAATCCGGATCGAGCTCTGGCTGCACCGAAGCAG
TCGCTGCAGCGGCAGTTGGCAGACGGCCTTCGCTTCTTGGTGAAGGACAAGTTCAAGTTGAATCAACCTAGCGGCCCGTC
TGATGGATGGCTGACCCAGGACGCACTCTGGCTGGTGAGCAAGCCTGCTGCCGATCAACTGAGAGCCTACCTGCTGGCCC
AGGGTATCGATGGGGTGCCCTCCTCTAACGCGCCGTTCTTCAGCATGCTCCAGGACCAAGCCGTCATCCAGACAAATGCC
GAGGACAAGGCCATTTGGACGGCCACGGTAGACAACGGTGCTGGATGGAGAAACAAGTTCACGCTACTCAAGATTGCTCC
AGCCTTGATCTGGACAGATGCTGCCGAGCGCCCCTCACCCTACAGCGGATCACTGGTCGTTGAAGATGGAACCGCCTCAA
CGGAAAAGCCGGAAACGACCTGTGAAATTCCCAACGGGCCGGCTGAACAGCAGCAAGCACCAGAAACGAAGATGATGCTC
CATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGATGATCAAGAAGA
AACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACGACTCGCCGGCTG
CCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGATTGCGCTCCTGAA
GCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATCTGGCATCGC
GGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCACGCCAGGAA
TTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGGAAGCTGGTG
CAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAAACCTGAACATCTGGACCATCAAGGTTTCTGGTCC
TCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGGACAACCCAA
GCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGTGGAATGACGCCGCAGCAGCTCACCGAGGAGTACATCTTCGCGCACG
ATCTCCGGGAAGCCAGCGCGAAGATCTACCGCGCCGCGACCAAGGCGCTGCTCAAGCACTTCGGTCCTACGGCAACCGTA
CAGGACGTGGACCACCGGGCTGTCCTGGGATGGCGACGCAAGGTACTGGAACAAGGCCTGTCCAAGCGGAGCTGGAACAC
GTACTCGAATCATCTGCGGACGATCTGGGGCTATGCCATCGAGCATGAGTTGGTGACGCACTCCCAAGTCAACCCGTTCA
GAAAGACCACCGTCATTCCCCCCAGGCGAGCAAGCAAAACCGTCGCCGCCGAAGCCATCCTGCTCGCCCGCAATTGGCTC
AACATGCAGGACGGCGCAGAGCGCTGCACCGGCGAACGCGCACGGATCACGCCCGCCTGGTTCTGGCTTTGCACGTTTGA
GGTCTTCTACTTCACCGGCATCCGGTTGAATGCGCTGTTGTGCATCCGCAAGCGCGACATCGACTGGGAAAATCAACTGA
TCCTCATCCGCGGCGAGACAGAGAAGACTCACAAAGAGTTCGTAGTGCCAATAACGGAGGGGCTTGTGCCTCACCTATCG
AGGCTCCTGCAAGAGGCCGATAGAGCCGGATTCGCCGATGACGACCAGTTGTTCAACGTCAACCGGTTCTCACCGCACTA
CAAGAGCAAGGTGATGAACTCCGACCAGGTCGAAGCCATGTACCGGAAGTTGACCGAGAAGGTTGGGGTGCGGATGACCC
CGCACCGTTTCCGGCACACCCTGGCCACCGACTTGATGAAGGCACCCGAGCGGAACATTCACCTCACGAAGTGCCTGCTC
AACCACTCGAATATCCAGACCACGATGAGCTACATCGAGGCCGACTACGATCACATGCGTGCCGTGCTGCATGCTAGAAG
CCTGGCCCAAGGCGCGCTGGAGAATGTCAGGAAGGTGGATTACAGCGGCTCCCCGCAAGCCTCTGCCAAACCGAAGCCAT
GCGGGCAACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAGGCCAGGACAGAGCCTGCAGAACCAAGGGAGCACACG
CCAGGGACAGGCATTCAGGGAGGTCCAACCGCGTGGGAAGCAGATGCGCTACCACAGCCACCTGACACCTTCGAACCAAG
CGTGCTGTTCACTCTGATGGCTCAAAACTTATCGAACCGTGCCGCCTCGGCATCCGCGGCTCCCGCTGCAACAAGCGGAT
CAGGCGGATGGGGATCTGCCGCCCGAAGCAATCTCGCCTAGCGATACCGGTACTGAGGGCCGGCTACCGGACGAAAGGTA
GCCGTGCCTTCCAGCAGATCGTTAGGCCTGTAGGAAAAATCTGGAATTACCGAGAGCGCCTGGATTCCAGCGCCGGCATG
CTGGCAGAGCCAGCGCAATTTCAAGGCCAATACCACAGTACCCTCTGTAATCGCTGATTACGTCGGGGGCGCATTGCTAC
GCCTGCAGAATGGTTTCAGGGCCTTAGAAACAGAAAAGCCCACCTAGAAAGGCGGGCTATTCCATATTGACATCACGTCA
ATGCGGGCCTAATGTTCGGCCCAGACGGCTGCTAGACAAGAACCGGCGTAACACCCCTTCCTAGCCTATGCAACTCGCCC
CGTAGAAAATGGTGGGTCGTGTAGGATTCGAACCTACGACCAATTGGTTAAAAGCCAACTGCTCTACCGACTGAGCTAAC
GACCCAAGTATGAGGTGGTCGGGGTAGAGAGATTCGAACTCCCGACATCCTGCTCCCAAAGCAGGCGCGCTACCGGACTG
CGCTATACCCCGATTGGAATTTGGCTCCGCGACCTGGACTCGAACCAGGGACCCAATGATTAACAGTCATTTGCTCTACC
GACTGAGCTATCGCGGAACGTCTTTCTTCCAACCCTGGACGCTTCCGGTGTTGCTGGATTCGCGTCTCAGAGGCGCGCCA
TTTTACGGATGCGCGCGGGCATGTCAACCCTCTGATCCAAAAAGTTTTTCTTCTTTTTCCACGAGCGACAAAACGGCCCT
TCCACTGCATGCGGCAGCGCTCTCGCGCCTACCGGACGCCCATGAAAAAGCCCCGCCGAAGCGGGGCTTTCCCTGTCCGC
CCCCGAAGAGGTCAGGCGAAGACGATCTCGTCGCCTTCCACCTTCGCCGAGATACTGGCACCCGGCGCGAATTTGCCGGC
CAGGATCAGTTGCGCCAGCGGGTTCTCGATCCAGCGCTGGATGGCCCGCTTCAGCGGGCGTGCGCCATAGACCGGGTCGA
AGCCGACGGCAATCAGCTTGTCCAGCGCCTCCTGGCTCAGTTCCAGGCTCAGCTCGCGCTCGGCCAGGCGCTTGCGCAGG
CGACCGAGCTGGATCTCGGCGATGCCGGCGATCTGCTCGCGAGCCAGCGGCTCGAACACCACCACTTCGTCGATCCGGTT
```

Fig. 2F

GATGAATTCCGGACGGAAGTGCGCATTGACCGCGTCCATCACTGCGGCACGTTGCGCCTCGCGGTCGCCGGCCAGCTCCT
GGATCTGCGCCGAACCGAGGTTGGAGGTCATCACCACCACGGTGTTGCGGAAGTCCACCGTACGCCCGTGACTGTCGGTC
AGGCGTCCGTCCTCGAGCACCTGGAGGAGAATGTTGAATACATCCGGATGGGCCTTCTCCACCTCGTCCAGCAGCACCAC
CGAGTAGGGCTTGCGGCGGATCGCCTCGGTCAGGTAGCCGCCTTCCTCGAAGCCGACGTAGCCCGGAGGCGCGCCGATCA
GGCGGGCCACCGAGTGTTTCTCCATGAACTCGGACATATCTATCCGCACCAGCGCCTCCTCGGTATCGAAGAGGAACTCG
GCCAGCGCCTTGCACAACTCGGTCTTGCCCACCCCGGTCGGGCCGAGGAAGAGGAACGAGCCGCTCGGCCGGTTCGGATC
GGCGAGGCCGGCGCGCGAACGGCGCACGGCGTTGGACACGGCGACTACCGCCTCGTCCTGGCCGATCACTCGCCGATGCA
GCTCCTGCTCCATGCGCAGCAGCTTCTCGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTGGAAACC
ACTTCGGCGATTTCCTCGTCGGTCACCTTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGTGCTGGTCGACCATCTGCAG
GCTGCGTTCCAGGTCCGGGATGGTCTGGTACTGGATGCGCGCCATGCTCTCGAGGTCGCCCTTGCGCCGCGCCGCCTCCA
TCTCCTGCTTGGCCTGCTCGATCTTCTGCTGGATCTGCGCCGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCC
TCGAGGTCGGCGTATTCGCGCTCGAGCTTGACGATATCCTCCTCCAGCTTGGCCAGGCGCTTCCTGGTGGCTTCGTCGTC
TTCCTTCTTCAGCGCCTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTCGAGACGATCCAGTTCCTCCGGCTTGGAGT
CGATCTCCATGCGGATGCGGCTGGCGGCCTCGTCGATCAGGTCGATGGCCTTGTCCGGCAGTTGCCGATCGGTGATGTAG
CGGTGCGACAGCTTGGCCGCGGCGATGATCGCGCCGTCGGTGATGCTCACCCCGTGGTGCACTTCATAGCGTTCCTTGAG
GCCACGGAGGATGGCGATGGTGTCTTCCTCGCTCGGTTCGTCCACCAGCACCTTCTGGAAGCGGCGCTCCAGCGCGGCAT
CCTTCTCGATGTACTGGCGATACTCGTCGAGGGTAGTAGCACCGACGCAGTGCAGCTCGCCGCGCGCCAGAGCCGGCTTG
AGCATGTTGCCGGCGTCCATGGCACCTTCCGCCTTGCCGGCGCCGACCATGGTGTGCAGTTCGTCGATGAACAGGATGAC
CCGGCCTTCCTGCTTGCCCAGTTCGTTGAGGACCGCCTTCAGGCGTTCCTCGAACTCGCCGCGGAACTTGGCACCGGCGA
TCAGCGCCCCCATGTCCAGGGCCAGCAGGCGCTTGTCCTTGAGGCCGTCCGGCACTTCGCCGTTGATGATGCGCTGGGCC
AGGCCCTCGACGATGGCGGTCTTGCCGACGCCGGGGTTCGCCGATCAGCACCGGGTTGTTCTTGGTCCGCCGCTGCAGGAC
CTGGATGGTCCGGCGGATCTCGTCGTCGCGACCGATCACCGGGTCGAGCTTGCCTTCCTCGGCGCGCTTGGTCATGTCGA
CGGTGTACTTGTCCAGCGCCTGGCGCGACTCCTCGACGTTCGGGTCGTTCACCGCTTCGCCGCCACGCAGGTTGGCCACG
GCATTCTCCAGCGCCTTGCGCGACACGCCCTGGCCGAGCAGCAGCTTGCCGAGCCTGGTGTTCTCGTCCATCGCGGCCAG
CAATACCAGCTCGCTGGAGATGAACTGGTCGCCCTTCTGCTGGGCCAGGCGGTCAGCCTGGTTGAGCAGGCGTGCGAGAT
CCTGGGACAGGTTCACGTCGCCGGTCGGGCTCTGGATCTTCGGCAGCGCGTCGAGTTCTTTGTTGAGGCCGCTGCGCAGG
GCGGCGATATCGAAGCCGACCTGCATCAGCAGGGGCTTGATCGAACCGCCTTGCTGCTCGAGCAGGGCGGAAAGCAGGTG
CACCGGCTCGATGGCCGGATGGTCATGGCCAACGGCCAGGGACTGGGCGTCGGAGAGCGCCAGTTGCAGCTTGCTGGTCA
AACGGTCTATTCGCATGGGTCGTCCTTCCTTCTATAGAGCGGGCCGGAACGATGGGTGTCCCTGATGAAGAAAAGCCCGC
CGAGATGACTCAGTAGATAAGGGCGATTTTCCGCGGTTCAAGCGACCGGACCGTGACATCGGTCAGTTGCCGCCGGATAA
CCTGCGCGGGCCTAGTCCTGGAGCCAGACCAGGCTGGCAAAACGGCCGGTACGCGACGAGCGGCGGTAGGAATAGAAGCG
CGCGGTATCGCTGAAGGTGCAGAAGCCGCCGCCATGCACGGCGGTGACGCCATGGGCGCCCAGGCGGATCCGCGCGAGTC
GGTAGATGTCGGCCATGAAGCGGCCCGGATTGGCGCTAGGTACGAAAGCCGAGCGCGCCTCGGCGTGCGCAGCGACGAAT
GCATCGCGGACCTCGCCGCCGACCTCGAAGGCCTGCGGGCCGATCGCCGGCCCCAGCCAGACCAGCAGTTCGTCGCCGGG
CACGCCCAGGCTGTCCACCGTCGCCTCCAGCACGCCCGCCGCCAGCCCGCGCCAGCCGGCATGGGCCGCGGCCACCCGGG
TGCCCGAGCGGTCGCAGAACAACGCCGGCAGGCAGTCGGCGGTCATGATCGTACAGGCGACGCCCGGCATCGCGCTCCAG
CTGGCGTCGGCCCTGAGCACCGGTTCGGGTCGGCCTCCACCACGTCACTCCGTGCACCTATTCCAACCAGCTCGGCCGGC
ATTCCAGACGCTCGGTCAGGCGTCGGCGGTTTTATTCCACGGCGCGCGGATCGTCGTAGACGTGGGCGCCAAGGTTCAGA
CTGTCGAAGGGTGCCTGGCTGACCCCGCCACTGCGCGTGGTCACGCAGGCCCGCACACGGGCCGGCGCCGGCCAGTCGGG
GGTCAGCCAGGCCGTTCAACCGACGAACGCCTCGCGATCCTGGCGCAACAGGCTGAGCAGCCAGAGGAATTCTTCCGGCAG
CGGCGATTCCCACTTCATGCGCACGCCGGTGGCCGGGTGATCCAGTTCGAGGAAGCGCGCGTGCAGCGCCTGCCGGGGGA
ATTCGCGAAGAGTCTGGACCAGGGTCTGGCTGGCCACCGGGGGAATCCTGAAGCGCCCACCGTAGACCGGATCGCCGACC
AGGGGATAGCCAATATGGCTCATGTGCACGCGGATCTGGTGGGTACGCCCGGTCTCCAGCTTGACCCGGGTATGGGTGTG
CGCACGGAAGCGTTCCAGCACGCGGTAATGGCTGACCGCCACCTTGCCGGCGTCGACCACCGCCATCTTCTGCCGCTGCA
CGCCATGCCGTCCGATCGGCGCATCGATGGTGCCGCCGGAGGTGATCACGCCGATCACGATCGCCTCGTAGATGCGGCTG
ACCGACCGTGCCTGCAGTTGCGCCACCAGCTTGGTGTGGGCCTCCAGCGTCTTGGCCACTACCATCAGGCCGGTCGTGTC
CTTGTCCAGGCGGTGGACGATCCCGGCGCGCGGCACATTGGCGATGTCCGGGACATGGTAGAGCAAGGCATTCAGCAGGG
TGCCGTCCTGATGGCCGGCAGCCGGATGGACCACCAGGCCGGCGGGCTTGTCAATCACCAGGATGTGCTCGTCCTCGTAG
ACGATTTCCAGCTCGATGTCCTGTGCGAGCCACTCGCCCTGGGCTTCCTGCTCGGCCTCCAGGACCAGTTGCGCGCCGCT

<div style="text-align: center">Fig. 2G</div>

GTGGACGATGTCGCGCGGGCGCAGCACGGCGCCGTCGACGGTCAGGCGACCGTCCTTGATCCAGCCGGCCAGACGGGAGC
GGGAGTGTTCGGGAAAAAGCTGGGCGGCGATCGGTCGAGACGCTGGCCACCCAGCTCGAACGGCACCTCGGCCGCGCGT
TGAATCATATCGGACATGAGTAGGAGACGATGCTCAGCGCGGCTTTTGGAATCGGCTACGCGCTGTGGTTAAATACGGGG
TCTTTGTCCCAGGGGGTGCCTGGGGCGCCAATCATAACAGACGGTTGAGGCCAAGCCGACCGTCCCAGGGACGCAAGCCG
CCATGCAAGTGAAACACCTGCTGCTGATCGCCATCCTCGCCCTCACCGCAGCCTGCTCCTCGAACAAGGAGACTGTCGAC
GAGAACCTGAGCGAGAGCCAGCTGTACCAGCAGGCGCAGGACGACCTCAACAACAAGAGCTACAACAGCGCCGTCACCAA
GCTGAAAGCCCTCGAATCGCGCTATCCCTTCGGCCGCTACGCCGAGCAGGCCCAGCTCGAGCTGATCTACGCCAACTACA
AGAACATGGAGCCCGAAGCCGCCCGCGCCGCCGCCGAACGCTTCATCCGCCTGCATCCGCAGCACCCCAACGTCGACTAC
GCCTACTACCTCAAAGGCCTGTCCTCCTTCGACCAGGACCGCGGCCTGCTGGCGCGCTTCCTGCCGCTGGACATGACCAA
GCGCGACCCGGGCGCCGCCCGCGACTCCTTCAACGAGTTCGCCCAGCTCACCAGCCGCTTCCCCAACAGCCGCTACGCCC
CGGACGCCAAGGCGCGCATGGTGTACCTGCGCAACCTGCTGGCGGCCTACGAAGTGCACGTCGGCCACTACTACCTGAAG
CGCCAGGCCTATGTCGCCGCCGCCAACCGCGGTCGCTACGTGGTGGAGAACTTCCAGGAAACCCCGGCCGTCGGCGATGG
CCTGGCGATCATGGTCGAAGCCTACCGTCGCCTGGGTCTCGACGACCTGGCCAGCACCAGCCTGGAAACCCTCAAGCTGA
ACTATCCGGATAACGCCAGCCTCAAGGATGGCGAGTTCGTCGCCCGCGAAAGCGAGGCCGACACCCGCTCCTGGCTGGCC
AAGGCCACCCTGGGCCTGATCGAAGGCGGCGAGCCGCCGCCGCACATGGAAACCCAGGCCGCCAAGGACGTGATCAAGCA
GTACGAGGATGCCGAGCGGGAGATCCCCGCCGAACTGAAGCCGGAAAACCAGGATCACAGCGCCGACGACGAGAAGCCGG
AGAGCGATGACGACGAGGACTCCGGTCGCTCCTGGTGGAGCTACATGACCTTCGGTCTCTTCGACTGATCGCACGAAACA
CCGAAGGGAGGCGCAGGCCTCCCTTCTTTTTGCCCGCCGCCATGCCTCTCCCAGCGCCAAACGCCGCACAGCCTGGACCT
TCCCGTCTGGGATCGAGCCGGCGGCTTGGCTAAACTGCAGCTTTCTCCAGCCTCCGAGATCACCATGGGCCTTTTCCGCC
TCCTGTTCTGGATCGCCCTGATCGCCATCGCGTTCTGGCTCTGGCGTCGCTTTACCCGTCCCACTCCGCGCCAGCAGCAA
CGTCCGCAGGACGAGCCGAGCGCATCGCCGATGGTCCGCTGCGCCCATTGCGGCGTCCACGTGCCGCAGGCCAACGCCCT
CGCCCACGAACAACGCTGGTATTGCAGCCAGGCGCACCTGCGCCAGGACCAGGGCGACCGTGCGCGCTGAACGGCTACGG
CTGAGCGAGGAGCAGGGGCAACGCATCCTCCGTCTGTACCACCTGTACCGCCTGACCATCGGCCTGGTACTGGTCCTGCT
GATCTCCAGCGAACTGGAAGATCAGGTCCTCAAGCTCGTCCACCCTGAACTGTTCCATGTCGGCAGTTGGTGCTACCTGG
TCTTCAACATCCTGGTCGCGCTGTTCCTGCCGCCGTCGCGGCAATTGCTGCCGATCTTCATCCTCGCGCTCACCGACGTG
CTGATGCTTTGCGGCCTGTTCTACGCAGGTGGCGGCGTACCCAGCGGCATCGGCAGCCTGCTGGTGGTGGCGGTGGCCAT
TGCCAACATCCTGCTGCGCGGGCGCATCGGCCTGGTCATCGCGGCGGCGGCCAGCCTCGGCCTGCTCTACCTGACCTTCT
TCCTCAGCCTGAGCAGTCCGGACGCCACCAACCACTACGTCCAGGCCGGCGGCCTCGGCACCCTGTGCTTCGCCGCCGCG
CTGGTGATCCAGGCTCTGGTGCGGCGCCAGGAGCAGACCGAAACGCTGGCCGAAGAACGCGCCGAGACGGTCGCCAACCT
GGAGGAACTCAACGCATTGATCCTGCAGCGCATGCGCACCGGCATCCTCGTGGTCGATAGCCGTCAGGCCATCCTCCTCG
CCAACCAGGCCGCCCTCGGCCTGCTCAGGCAGGACGACGTGCAGGGCGCCAGCCTCGGCCGCCACAGCCCGATGCTGATG
CACTGCATGAAGCAATGGCGCCTGAATCCCAGCCTCCGTCCGCCGACGCTCAAGGTGGTGCCGGATGGCCCGACGGTGCA
ACCCAGCTTTATCAGCCTCAACCGCGAAGACGACCAGCACGTGCTGATCTTCCTCGAAGACATTTCGCAGATCGCCCAGC
AGGCGCAGCAGATGAAGCTGGCCGGTCTTGGCCGCCTGACCGCCGGCATCGCCCATGAGATCCGCAACCCGCTGGGCGCG
ATCAGCCACGCCGCCCAACTGCTGCAGGAGTCAGAGGAACTGGATGCCCCGGACCGACGCCTGACGCAGATCATCCAGGA
CCAGTCGAAGCGGATGAACCTGGTCATCGAGAACGTCCTGCAGCTCTCCCGTCGCCGCCAGGCCGAACCGCAGCAGCTCG
ACCTGAAGGAGTGGCTTCAGCGGTTCGTCGACGAATACCCCGGCAGGCTGCGCAACGACAGCCAACTGCACCTGCAGCTC
GGTGCCGGCGACATCCAGACCCGCATGGACCCACACCAGTTGAACCAGGTGCTGAGCAACCTGGTGCAGAACGGTCTTCG
CTACAGCGCCCAGGCGCACGGGCGCGGCCAGGTCTGGCTGAGCCTCGCGCGCGACCCGGAGAGCGACCTGCCGGTGCTGG
AAGTCATCGACGACGGTCCCGGCGTACCGGCGGACAAACTGAACAACCTGTTCGAACCCTTCTTTACTACAGAAAGCAAA
GGCACCGGCCTGGGCCTCTATCTCTCCCGCGAACTCTGCGAGAGCAACCAGGCACGGATCGACTACCGCAATCGCGAGGA
AGGCGGCGGCTGCTTCCGCATCACCTTCGCCCACCCGCGCAAACTCAGCTGACGGAAGCCGCACGCATGAGCCGACAAAA
AGCCCTGATCGTCGACGATGAACCGGATATCCGCGAACTGCTGGAAATCACTCTCGGCCGCATGAAGCTGGACACCCGCA
GCGCCCGCAACGTCAAGGAAGCCGCGAGTTGCTGGCCCGCGAGCCGTTCGACCTGTGCCTCACCGACATGCGCCTGCCGG
ACGGCAGCGGCCTCGATCTGGTCCAGTACATCCAGCAGCGCCATCCACAGACCCCGGTGGCCATGATCACCGCGTACGGC
AGCCTGGACACCGCGATCCAGGCGCTCAAGGCCGGTGCCTTCGACTTCCTCACCAAACCGGTCGACCTCGGTCGCTTGCG
GGAGCTGGTGGCAACCGCCCTACGCTTGCGCAACCCGGAAGCCGAGGAAGCGCCGGTGGACAACCGCCTGCTCGGCGAGT
CGCCGCCGATGCGCGCCCTGCGCAACCAGATCGGCAAGCTGGCGCGCAGCCAGGCGCCGGTCTACATCAGTGGCGAGTCC
GGCAGCGGCAAGGAACTGGTGGCGCGCGCCTGATCCACGAGCAGGGGCCACGTATCGAGCGGCCGTTCGTGCCGGTGAACTG
CGGCGCGATTCCCTCCGAGCTGATGGAAAGCGAGTTCTTCGGCCACAAGAAAGGCAGCTTCACTGGCGCTATCGAAGACA

Fig. 2H

```
AGCAGGGCCTGTTCCAGGCCGCCAGCGGTGGCACCCTGTTCCTCGACGAAGTCGCCGACCTGCCGATGGCCATGCAGGTC
AAACTGCTCCGGGCGATCCAGGAAAAGGCCGTGCGCGCGGTCGGCGGCCAGCAGGAGGTCGCCGTCGCACGTGCGCATCC
TCTGCGCCACCCACAAGGACCTCGCCGCCGAAGTCGGCGCCGGGCGCTTCCGCCAGGACCTCTACTACCGCCTCAACGTC
ATCGAGCTGCGCGTACACCGCTGCGCGAACGCCGCGAGGACATCCCGCTGCTCGCCGAACGCATCCTCAAGCGCCTGGCC
GGCGACACCGGCCTGCCGGCCGCCAGGCTGACCGGCGACGCACAGGAGAAGCTGAAGAACTACCGCTTCCCGGGCAACGT
CCGCGAGCTGGAAAACATGCTGGAGCGCGCCTATACCCTGTGCGAAGACGACCAGATCCAGCCTCACGACCTGCGCCTGG
CCGATGCGCCGGGTGCCAGCCAGGAAGGCGCCGCGAGCCTGAGCGAAATCGACAACCTCGAGGACTACCTGGAAGACATC
GAGCGCAAGCTGATCATGCAGGCACTCGAGGAGACCCGCTGGAACCGCACCGCCGCGGCCCAGCGCCTGGGCCTGACGTT
CCGCTCGATGCGCTACCGCCTGAAAAAGCTGGGCATCGACTGAAAGTGAAAAGGCCTGTCCGAAGACAGGCCTTTTGGTT
TTCGCTCCTCAGAGGCGACCAGCCGGGGCGTAGGGGGCCGGGTCGATGATCGGTTCCCGCCCGCTCATGAGATCCGCCAG
CAGACGGCACGACGCCGGTGCCAGGACCAGCCCGTTGCGGTAGTGCCCGGTATTCAGCCAGAGCCCGTCGAAGCCAGGCA
CCGGACCGATATAGGGGATGCCTTCGGGAGAGCCCGGGCGCAACCCTGCCCAGTGGGCCACCGGCTGCATGTCCGCCAGT
TCCGGCAACAGTTCTGCCGCAGACGCCCTGAGGCTTTCCAGCGCCTCGTCGGTCGGCGTCTTGTCGAAGCCCGAATGTTC
CAAGGTGCTGCCGATCAGGATGTGGCCGTCGCGCCGCGGAATCGCGTAGCGCCCCTTGGCCAGCACCATGCGCGGCAGGA
AATCCGCCGCGCACTTGTAGAGGATCATCTGACCTTTCACCGGTACCACGGGCAGTTCCAGGCCAAGCGGCTTCAACAAC
TCGCCGCTCCAGGCGCCTGCCGCCAGCAGCACCTTGTCGCCACGGATCTCGCCACGCGAGGTCGCCACGCCGACCACTCG
ATCGCCGTCGCGCAACCAGCCGCGCACCTCCGTCTGTTCATGCAACTCGAGATTGGCGAATTGTTGCAGGGATGCCCGCA
ATGAGCGCGCCAGGCGAGGATTGCGCACATTGGCCACGCCCGACATGTAGACCGCCCGCTGGAAGCCTGCGCCCAGCCCG
GGCACCGCCGCGTAGGCCTCCTCGATCGGCACTTCCTTCAACGGCCGGGTGTGGTTGCGTGCCCACTGCAGTGCCTCGGT
CTGGTCGTCCAGGTCCAGCCAGTACAGGCCAACGGTATGGACCTCGGGATCGAGCCCGGTCTCGTCGAGCAAACGCTGCC
CCAGGGCCGGGTAGAAGTCCTGCGACCAGTGCGCCAGGGCGGTCACCGCCGGGCTGTAGCGCCACGGATAGAGCGGCGAG
ACGATCCCGCCTCCCGCCCAGGATGCCTCACGCCCACTCTCGCCCCGCTCCACCAGGGTCACCCGCAGTCCGGCGAGCGC
CAGCTCCCGGGCGGTCAACAGGCCGATGACGCCAGCGCCTACCACTACTACATCTCTACTCACCACAGGGCTCCTACCGA
TTTGCCAGGAACAGAGAAATATCACTCAAAGGGATCAGATGCTGACGAATTGCCTGCTTCAACGAACTCAGTCGAATCTA
GTCCCGGTGAAAAGCCCATCATACCCGCAGAGGTATTCATCCCATGAAATCGAGTGGTTTGAATTTGGTGGAACTATCGA
TAGTCCTATCGATCCTTGCGATAGGCGTGACAATTGCGCTGCCCACCCTCCCCGACAGAATGAAGCGGGACATTAGCCGT
GATATTGGTGACAGCCTGACTAGTCATGTGATGGCTGCGCGGGCTAGCAGCATACAGAACGGCGTGATCATCGAGGTGTG
CGGTAGCGGTGACGGCAGTACCTGCAGCGAGGAATGGCATCTCGGCTGGTTCAGCCGTAACGACAGGAGCCAACAGATAC
TGGCCCGGCATGAAAATACGAGTCGCACCGATATTCATTGGCGGGGCTTCGACAAGCGACTGCGCTACCTGCCTAATGGC
ACCAGCCCTACAGGTAACGGGCGTTTCTTCGAATGTAAGGACGATCGCATCGAGTGGCAATTGGTGCTCAATCGGCAAGG
CCGCCTCAGGGTGGCGGGAAAGAGCGAAAATAAAAAGCTCTCTTACCTGTGCTCCAGGCGGTGAGAGAACTGTTTCACAT
ACCGTTTGCCAGTCATCCCACTCTCCGCTCCGGCTGTCTCTGCTACAGGGACAATGCGCTCTCCACTAGGCAAGATTATC
TGGCCCTTTTCCTTGTGGAGTACTGCATGCGCTCTATTTGTCGCAGCGCCGGCTTTTCCCTGATCGAGTTGATGATGGTG
TTGGTTCTGGTCGCCATATTCGCCAGCATTGCCGTACCCAGTTTCAACGCCTTGATCGAGCGCAACCGAATCCAGACTGC
CAGCGAGGAACTCTACAGCCTGCTTCAGTACGCTCGCAGCGAAGCTGTAAACCGTCATGCCAATGTGAGCATCAGGGCGA
CGCAGAACAATGACTGGGCAAAAGGCCTGGAAATCATCAGCGGCGCGACCACCGTGCAAAAGCACCAAGGTTTCCAGCAG
GTCTCGCTATCCGCCAGCAGTGCGACTGCGGAGCTGACCTTCAACGCTACCGGCACACTTAGCAACCAGGCTGCAAACAT
TGACATAAAGGTCTGCTTCGCCGGTGACAAAAGTACAGGACGTCTGCTTACCGTTCAGCCCAGTGGACGCGTGATCCTGT
ACCCATCTTCAAAGCAACCGGACAGCTGTAACTGAGGAAAGCCCATGTCTCGAGAAACGGGTTTCAGCATGATCGAAGTA
CTGGTTGCTCTGGTGCTGATCAGCATTGGCGTACTGGGCATGGTTGCCATGCAAGGGCGCACGATCCAGTACACGCAGGA
GTCGGTACAACGCAATGCCGCAGCAATGCTTGCTAGCGACCTGATGGAAATAATGCGTGCGGACCCAGATGCCGTACTCA
ATCTACGCGCCCAACTACGCGAAGACTCGGTCTACTACAAGGCCAAGGGCAGCGACTTTCCCGCAGCCCCAGCGCGCTGC
GCGCCATTGCCAGCAGATGCTAAGGAACGTCTCGGCTGCTGGGCCCAACAGGCCTCGAAAGACTTGCCGGGAGCCTCCGC
ACTCTTGAATAGCCAATTCTACATTTGTCGCAGCCCAACCCCGGGTACCTGCGACAACACCAAAGGCTCGGCCATCGAAA
TCCAGGTTGCCTGGCGAGCCATGGATGGAGCGTGTTTCAACGCCTCTGACTCCACCTTGTGCACCTACAGCGTCCGCTCC
GAATTGTGAGAACAAGCATGCTCTTCAGCAAAATGCAGAAAGGCCTATCGATGGTAGAACTGCTCGTGGCACTCGCTATA
AGCAGCTTCCTGATCCTGGGGATCAGCCAGATCTACATCGACAACAAACGCAACTATCTTTTCCAGCAAGGCCAGGCCGG
CAACCAGGAAAATAGCCGCTTCGTTCTTATGCTGCTGCAGCAACAACTGGATAAGACAGCCTATCGTCGCCTTCACGACG
ACAACATGGAGAATGCTTTCAAATCCGCGACATTCAATGGCTGTCGTGCATTTGTGGCTGGCGAGACTATCGCTGCGGCA
ACTGCCCTCAAGGCGGGTGAGTACGGTGTCTGCTTGCGCTATCAACCCGCCTACAAAGGGGAGCATGATTGCCTCGGTAA
```

Fig. 2I

```
TGAAATTACCGGAGTTCCGGAAAAGCCCTTCACAAATACTCCCCCTGTCGTCGTTCGCCTGGTCTACCTACCGAGCGCCG
GTACCCTGAGTTGCAGTCGTCCCGATATCGCCCAGTCGAAATCGGGAGAATTGGTCAGTGGTCTCACAGACTTCCGCTTG
GAAGCGGGGGTCGGGCCAGCAGATCGTAGCGAACGCAAAGTATCCAGCTTCGTCGCACTACAGGATGTCGCCGGTCGTCC
TATCCGAGCATTGCGCTTCTCAATCCTGGCAGGCAGCGACAATACAAGCCTGCGCACAGGAGATGATAGCCAGGCACGCG
ATCGCTGGATCGTCCTTTATCCCGAGAGCAAAAGCGCCATCGAGGCCGCAGACAAAGGCCAGATTTACCAAATAGCGCGT
GGTAACCAAACCATCAGGAATCTCATGCCATGACCCTGCGCCATACCTCTCGACAGCAGGGATCCACGTTGTTGATCTCG
CTGGTTATCTTGTTGATGATCACGCTCCTCGCCGTTTCCAACATGCGCGAGGTGTCACTGGAAAGCCGTATCACCGGCAA
TCTCATCGAACAGAAGCGCCTGCGCAATGCGGGCGAAGCTGGGCTACGCGAAGGTGAACGACGCTTTTTCAATACCATCA
AGCCCCCAGAGGTCGGCAGCGGATGCGCCGATAGCAATGTCAAACGGCCTTGCATACTGAACCTGAGTGCCCTCTCCGTA
CCCCGAGATGACGTGCACAACAATCCGGTGGCAGCCCTGAACGGCAAGACAGATAACGCCAATTCACGTGTCTGGATGCC
CTACCGAGGCAGCGATCTGAATAACCCTACGCAGATCGACAAAGACCGCGCAGTCACCTGGCAGACCATCACGGTGCCCG
CTGGCGAACAGAACAACGAAGCGGAAAATCCCGAGTACGGCAACATGATGCGCGGGGTCGGCACGTTCTACTACGAAACC
AACAGCCGCGCCCTCAACAAGGCGGGCGGAGAGACTGTTCTACAGGCCGTTCATGCACGCCTGTATACCAACTGACTGGA
GCCAGCGCATGATCCACCAGATTACCCGCGCAGGAAAAAGCCTGCTGGCTGCAGGTTGCACCCTGAGCATCCTGTTCGCC
TCTGACAGTTATGCCGCCACGGCCCTGAATGTCAGCCAGCAACCCCTGTTCCTAACCCAGGGCGTTGCTCCCAACCTGCT
GTTCACTCTAGATGACTCAGGCAGTATGGCCTGGGCTTACGTGCCCGACGGTATTAGCGGGAATAGCGGCAGAGCGGGAC
GTTCCAGCGATTACAACGCACTGTACTACAACCCCGATTATGCTTACCAAGTGCCCAAGAAATTGACACTGTCAGGCGAT
CAGATCATCGTTTCCGACTATCCAGTGCCACGCTTCACAGCAGCCTGGCAGGATGGCTACGCCCAAGGCTCCACCACCAA
CCTGAGCAATAACTATCGCCCTCAATGGGGAACCGGCTGGCTTGGTTGCATCGATAGCAGCTGCAATACCGGGAGAGCTT
ATTACTATACTTATAAGGTAAGCGCTAGCTGCCCTGCACAGCCGGTGAGCAGCTCCAACTCCTGTTATACCTACAATGCT
CTTCCTACCAGTCAGGAAAGCAACTTTGCGATATGGTACTCCTACTATCGCAACCGCATCCTGGCCACAAAGACCGCTGC
CAACCTGGCCTTTTACAGCCTGCCGGAAAACGTGCGTCTCACTTGGGGGGCCCTGAACACCTGTAGCATCGGCGCCAACA
GCAGAAGCTGCCAAAACAATGCCCTGCTCCAATTCAACAAGCAGCACAAAATCAATTTCTTCAATTGGCTGGCGAACAGC
CCGGCCAGCGGCGGTACTCCTCTGCATGCGGCTCTTGACCGAGCCGGACGCTTCTTGCAAACCAACGGCACAGCTTATAC
CACCGAAGACGGAAAGACATATTCCTGCCGGGCCAGCTATCACATCATGATGACCGACGGTATCTGGAACGGTCGGAACG
TCACCCCCGGCAATCTCGACAACCAGAACCAGACCTTTCCTGATAGCACCCTCTATAGGCCACAGCCCCCTTATGCCGAC
AGCAATGCCAGCTCATTGGCTGACCTGGCTTTCAAATACTGGACCACAGACTTACGTCCCAGCATCGACAATGACCTGAA
GCCTTTCATGGCCTACAAGAGTGGGGACGATTCCAAGGATTACTGGGACCCTCGCAACAACCCAGCCACTTGGCAACACA
TGGTCAACTTTACCGTTGGCCTAGGTCTTTCCTATTCGCTCACATTGAACTCTGCACCAACTTGGACAGGCAGCACCTTT
GGCAACTACGAGGAGTTGATGGCTGGAAGCAAGGCTTGGCCCAGCGTCGATAACGACGCCGCACCCGGTAACGTCTACGA
CCTCTGGCATGCAGCTATCAACTCTCGTGGAGACTTCTTTAGCGCGGAATCACCGGACTCTCTGGTTCAGGCTTTCAATA
AGATCCTGACACGGATTTCCGAGCGCAACACCTCCTCCTCCAAACCAGCAATGACTTCCGCGCTGCAGGATGACGGAACC
GGCGACAAGCTGATCCGCTACAGCTACCAGTCCAGCTTTGCCAGTGACAAGAACTGGGCGGGCGACCTTATACGTTACAA
GGTGGAGTCGACTTCCACCGGTTCGACCAAAAACCCAGGAATGGAGCGCCGGCGCACTGCTGGACAACCGAGCTCCCGCTA
CCCGTAATATTTACATCGCCAGCAATAGCGGAACCAACCGCCTTAAGCCTTTCACATGGAGCAATATTGAGGGAAGTCAG
TTAGCCACTTGGCTGAACCGCAACCCGGACAAGGACAATCAGGCCGACACCAAAGGAGCACAGCGGGTCGACTTCATCCG
TGGCCAGCAGAATATGGATGGATTCCGGCAACGACAGGCGGTGTTAGGGGACATCGTGCACTCGTCTCCAGCCGTGGTCG
GACCGGCCCAATACCTCACTTATCTGGCCAACCCCATCGAACCCAGCGGCGACTACGGCACATTCAAGACAGAGGCAGAC
CAGCGCAGCCCTAGAGTTTATGTTGGATCCAACGATGGCATGTTGCATGGTTTCAACATCAAAACCGGCGTGGAAGAGTT
CGCTTTCATCCCTACAGCAGTATTCGAAAAGCTTAACAAGCTTACCGGCATCAGCTACCAGGGCGGTGCCCACCAATATT
TCGTCGACGCTACACCGGTCGTCAGCGATGCCTTTTTCGATGGAGCTTGGCACACTGTTCTGATCGGAACGCTTGGTGCT
GGAGGTCGCGGCCTGTTCGCACTCGATGTAACCAAGCCGGACGATGTCAAGCTGCTTTGGGAATACGATAGCAGTACCGA
CTCGGACCTTGGTTACACCTTCTCCAAACCTACCGTAGCCAGACTGCACAGCGGACAATGGGCAGTAGTTACCGGCAACG
GCTATGGAAGCGATAATGACAAGGCAGCTTTACTGCTGATTGATTTGAAAAAGGGAACGCTGATCAAGAAGCTGGAAGTC
CAAAGCGAGCGCGGAATAGCCAATGGCCTATCGACGCCTCGCCTGGCTGATAACAACAGCGATGGCATTGCTGACTACGC
CTATGCTGGCGATCTGCAGGGAAATATCTGGCGCTTCGATTTGATCGGCAATACCCGCAACGACGACCCAGACACAAATA
CCTCTATCAATCCCTTCAAGCCCGGAGATGTAGATCCTTCTGCTTTCAGAGTATCGTTCAGCGGCGCCCCGCTTTTCCGT
```

Fig. 2J

GCTCGCGCCGACAACAATACTCGTCAGCCCATCACGGCTCCGCCTACCTTGGTACGCCATCCTAGCCGTAAGGGCTACAT
CGTCATCGTAGGTACAGGAAAATACTTCGAGGACGATGACGCTCAGGCCGATACCAGCCGAGCCATGACGCTCTATGGTA
TCTGGGATCGCCAGACCAAGGGCGAAAGCGCAAACAGTACCCCAACCATCGACCGCAACGCCCTCACAGCCCAAACCATG
ACAACAGAGGCGAACTCCACATTCGGTAGCGTGAACAGGAATATTCGGCTTATTAGCCAAAACCCGGTGAAGTGGTACAA
AGACGGAGCAACCGGTACCGCGAACTCGGATGTGGCTAGCTATGGCTGGCGACTGAATCTGGAGGTCAATAGCAGCAAGA
AAGGCGAAATGATGATCGAAGATATGTTCGCTGCCGGCCAAGTGCTTCTATTGCAGACCTTGACACCGAACGACGACCCT
TGTGACAGCGGCTCTACCAGCTGGACCTACGGCCTCAATCCATATACTGGCGGACGTACCAGTTTCACCGTCTTCGATCT
CAAACGTGCGGGTATAGTGGACTCTGGCTCGGATTACAACGGCTCGGTCGTATCCGCCTTCCAACAGGATGGACTAGGTG
GCTTGGCCATTACCCAGAACGAACAGCGTCAATCCGAGGCTTGCACTGGTGATGAGTGCATCATCTTCAACCCCAGCGAC
AAGAGTAACGGACGACAAACCTGGCGGGTCGTCGAGGAGAAATGAACATGAACCCCTTACGTCTTCTCGCCACAGCTCTT
GCAGCTCTAGCTCTGGCTTGCCCAACCTTTGCCTTGAGTGCCACGAATACGTTCGAGAATGTGGGCGTGGTCGAGGATGT
TCATCCTGCCGCCGGTCTGGTAGTAGTCGATGGGCAGACATATCGCTTGCCCAACCGTGTCCAACAACAGGACTCGCCGG
TCATATTCTTGGTACGTCAGGGACAGACAGTGTCTTTCTCCGGCAAACTCACCAGCGACCTGCCAGAAATCGAGTCGTTC
TACATTATCAAGCAGGCCCCTCTCGTTCCCTTCGGATCGGAGCAGCAACAATGAAGTCGAACAGAGGCTTCACTCTCATC
GAGTTGATGATCGTCGTAGTAATCATCGCTATTCTTGCTGGTATCGCCTACCCCAGCTACGACGAATACGTGAAGCGCGG
GAATCGCACCGAAGGACAGGCATTACTCAGCGAAGCAGCCGCTACTCAAGAGCGCTATTTTTCACAGAACAATACTTATA
TCACTACCCAAGCCGACATCGGCAAGCTGCATATGCGCAACACATCGGGCACCACAGTGAAGTCCTCCACAGGCAAATAC
AGCCTTACCGTCGATACGGTAGCCAACGACGGAGGTTATCGCCTTATCGCTAACCAGGCATTCAACGATCTTGATTGTGG
CAACCTGACCTTGACCGCCAACGGCGAGAAAGGCCGGACTGGAAGCAAGAAGAGCGTTGCAGAATGCTGGCGCTAAAGCG
CCGAGACAAGAAAAAGGCAAAGCCCGGCATAAGCCGGGCTTTTTCAGGTGCGCAAAAATTCCGATTACAAAGCCTTGACC
CGCAGTTCCTTGGGCATCGAGAAGGTAATGTTCTCCTCCCGTCCCTCCAGTTCCTGCTCTTCCGACGCCCCCCACTCACG
TAGCTGGGCGATCACTCCGCGCACCAGCACTTCCGGCGCGGAGGCGCCTGCGGTGATTCCGATGCGACGCACACCGTCGA
ACCAGCCGCGTTGCATGTCCTCGGCGCCGTCGATCAGGTAGGCCGGCGTGCCCATGCGCTCGGCGAGTTCGCGCAGGCGG
TTGGAGTTGGAACTGTTGGGGCTGCCCACCACCAGGACCATGTCGCACTGGTCGGCCAGTTCCTTCACGGCATCCTGGCG
GTTCTGGGTGGCATAGCAGATGTCGTTCTTGCGCGGCCCCTGGATCTGCGGGAACTTGGCGCGCAGGGCATCGATGACCT
TCGAGGTGTCGTCCATCGACAGGGTGGTCTGGGTCACGTAGTGCAGGGCTTCGGGCTTGCGCACCTCCAGCGCGGCGACG
TCGGCCTCGTCCTCCACCAGGTAGATGGCACCGCCGTTGCTGGCATCGTACTGGCCCATGGTGCCTTCCACCTCGGGGTG
GCCTTCATGCCCGATCAGCACGCATTCGTGGCCGTCGCGGCTGTAGCGCACCACTTCCATGTGCACCTTGGTCACCAGCG
GGCAGGTCGCGTCGAAAACCTTCAGGCCGCGCCCCTCGGCTTCCTTGCGGACCGCCTGGGAAACGCCGTGGGCGCTGAAG
ATGACGATGACGTTGTCCGGCACCTGATCGAGTTCCTCGACGAAGATGGCGCCGCGCTGGCGCAGGTTGTCCACGACGAA
CTTGTTGTGCACCACCTCGTGACGCACGTAGATCGGCGGGCCGAAGACATCGAGGGCACGGTTGACGATCTCGATGGCGC
GATCCACGCCGGCGCAGAAGCCGCGGGGATTGGCGAGTTTGATTTGCATGGCGGTCTCGTGGGCGACGCGGTGATTGGAC
GAATGAACCTTGCTACCGCCCTCCCCGCTTGGGAAGGGCGCAGCGACCGACGGTTCAGGCCGGCTGGACGTCGAT

Fig. 2K

>ORF2 (SEQ ID NO:2)
TCGCCGATCCAATGCCAAGGAGTACCTGGGCAATCAGAGCCTACTCACGGCTGCCGGGGCCGGCATTGCCAAGCTCCTGG
ACGCCGACGAGAACAACACCAGTACCGTCTTCAGCGGCAACGGCACCAGCTTCGGGACGACCGGAACCAACAGCAACTCG
GCCCTCAACAGCATCCTCTCCGGCGGCGTCAGCGACATCCGGCAGTGGATGAACAAGTTGTACGGGGAGGCCTTCGCCGC
CGTCTACGTGCAGCCAGGTGCGCGGGTCGCAGTGCATCTCGATCAGCAACTGGCGATCGACTATGA

>ORF3 (SEQ ID NO:4)
CGCCGATCCAATGCCAAGGAGTACCTGGGCAATCAGAGCCTACTCACGGCTGCCGGGGCCGGCATTGCCAAGCTCCTGGA
CGCCGACGAGAACAACACCAGTACCGTCTTCAGCGGCAACGGCACCAGCTTCGGGACGACCGGAACCAACAGCAACTCGG
CCCTCAACAGCATCCTCTCCGGCGGCGTCAGCGACATCCGGCAGTGGATGAACAAGTTGTACGGGGAGGCCTTCGCCGCC
GTCTACGTGCAGCCAGGTGCGCGGGTCGCAGTGCATCTCGATCAGCAACTGGCGATCGACTATGAACTCAAGGGCCGCAA
GGTCGATTACAGCTCTGGAGCCGCTCATGCAACAGCAGACTTGGACTAA

>ORF602c (SEQ ID NO:6)
TCGGCCTGGTCATTCGCGGAGGCATCCTGCTGCGGATCGATTGGGCGCCGCAGCGTTTGCCTGGCATCGAGGAGCAGCCG
GCCACGGCTGTTGCCTATCGAGCTGGTCGCACCTCGCTCCCAGACGTCGAGCATGTTGGCCTCGCCGTGGGGCAGCATCT
CCTCCTTGCTGGTGGAGCACGCTGCCAGGGTCAGCGCGCAGGCCAGGCCGGCGCAGAGACGAAGAAGGGGGTTAGTCCAA
GTCTGCTGTTGCATGAGCGGCTCCAGAGCTGTAATCGACCTTGCGGCCCTTGAGTTCATAGTCGATCGCCAGTTGCTGAT
CGAGATGCACTGCGACCCGCGCACCTGGCTGCACGTAGACGGCGGCGAAGGCCTCCCCGTACAACTTGTTCATCCACTGC
CGGATGTCGCTGACGCCGCCGGAGAGGATGCTGTTGAGGGCCGAGTTGCTGTTGGTTCCGGTCGTCCCGAAGCTGGTGCC
GTTGCCGCTGAAGACGGTACTGGTGTTGTTCTCGTCGGCGTCCAGGAGCTTGGCAATGCCGGCCCCGGCAGCCGTGAGTA
G

>ORF214 (SEQ ID NO:8)
ACAAGTTGTACGGGGAGGCCTTCGCCGCCGTCTACGTGCAGCCAGGTGCGCGGGTCGCAGTGCATCTCGATCAGCAACTG
GCGATCGACTATGAACTCAAGGGCCGCAAGGTCGATTACAGCTCTGGAGCCGCTCATGCAACAGCAGACTTGGACTAACC
CCCTTCTTCGTCTCTGCGCCGGCCTGGCCTGCGCGCTGACCCTGGCAGCGTGCTCCACCAGCAAGGAGGAGATGCTGCCC
CACGGCGAGGCCAACATGCTCGACGTCTGGGAGCGAGGTGCGACCAGCTCGATAGGCAACAGCCGTGGCCGGCTGCTCCT
CGATGCCAGGCAAACGCTGCGGCGCCCAATCGATCCGCAGCAGGATGCCTCCGCGAATGACCAGGCCGACTACACCCGCA
CGGCCAGCAACGAGATCCACAGTCAGTTCAAACGACTGCCCAATCCCGACCTGGTGATGTATGTGTTCCCGCACCTGGCC
GGCAGCGATCCCGCCCCGGTACCGGGCTACACCACCGTGTTCCCCTTCTACCAGCGAGTCCAGTACGCCATGCCGGGCGA
ACGCACGGAGGACTATTGA

>ORF1242c (SEQ ID NO:10)
TCTCGTCCTGGGCGTAGAACTGGACAATCCAGGGTGAGGTTTCGTGCTCGTCGAAGGAGTTCTGCAGGGCTTCTTTCAAT
GCGTCCCGGGCGTTCTGCATCCAATTGGGATCGCGGCCCTCGGTGCCCAAGGGCACCAGTTCGAAGAATGCGGCGCGCGA
ACGCCCATCCTCCAGGAGCATCACTTGCTCGTCGGGCAGGTACTCAGCCCAGGGCAGCAGGTCTACGAACGATGGGTCGT
GATCGTAGAGGCGTGACGCTTCGGCCTGCGTGGCGCCATTCTTGCTCCCGGTGTTGGGCAGAGGAATACCCATGGCGGCC
AGTCGCGCCAGATAGCGCTCAGTCGCTTCTTCCGCGGCCGCTACGTCCAGCGCTCCTGAATCTTCGGGAGCGTCTGCCGG
TACCGACTGAGGCTGTGTGCGACCGCGCAGAAGGGTTTGAAAAAAGCCCATCAATAGTCCTCCGTGCGTTCGCCCGGCAT
GGCGTACTGGACTCGCTGGTAGAAGGGGAACACGGTGGTGTAGCCCGGTACCGGGGCGGGATCGCTGCCGGCCAGGTGCG
GGAACACATACATCACCAGGTCGGGATTGGGCAGTCGTTTGAACTGACTGTGGATCTCGTTGCTGGCCGTGCGGGTGTAG
TCGGCCTGGTCATTCGCGGAGGCATCCTGCTGCGGATCGATTGGGCGCCGCAGCGTTTGCCTGGCATCGAGGAGCAGCCG
GCCACGGCTGTTGCCTATCGAGCTGGTCGCACCTCGCTCCCAGACGTCGAGCATGTTGGCCTCGCCGTGGGGCAGCATCT
CCTCCTTGCTGGTGGAGCACGCTGCCAGGGTCAGCGCGCAGGCCAGGCCGGCGCAGAGACGAAGAAGGGGGTTAGTCCAA
GTCTGCTGTTGCATGAGCGGCTCCAGAGCTGTAATCGACCTTGCGGCCCTTGA

Fig. 3-1

>ORF594 (SEQ ID NO:12)
CCAGGCCGACTACACCCGCACGGCCAGCAACGAGATCCACAGTCAGTTCAAACGACTGCCCAATCCCGACCTGGTGATGT
ATGTGTTCCCGCACCTGGCCGGCAGCGATCCCGCCCCGGTACCGGGCTACACCACCGTGTTCCCCTTCTACCAGCGAGTC
CAGTACGCCATGCCGGGCGAACGCACGGAGGACTATTGATGGGCTTTTTTTCAAACCCTTCTGCGCGGTCGCACACAGCCT
CAGTCGGTACCGGCAGACGCTCCCGAAGATTCAGGAGCGCTGGACGTAGCGGCCGCGGAAGAAGCGACTGAGCGCTATCT
GGCGCGACTGGCCGCCATGGGTATTCCTCTGCCCAACACCGGGAGCAAGAATGGCGCCACGCAGGCCGAAGCGTCACGCC
TCTACGATCACGACCCATCGTTCGTAGACCTGCTGCCCTGGGCTGAGTACCTGCCCGACGAGCAAGTGATGCTCCTGGAG
GATGGGCGTTCGCGCGCCGCATTCTTCGAACTGGTGCCCTTGGGCACCGAGGGCCGCGATCCCAATTGGATGCAGAACGC
CCGGGACGCATTGAAAGAAGCCCTGCAGAACTCCTTCGACGAGCACGAAACCTCACCCTGGATTGTCCAGTTCTACGCCC
AGGACGAGATCAGCTGGGACAATTTCCAGGAGCAGTTGAGGCAGTACGTCCATCCTCGAGCGCGAGGATCGGCCTTCAGC
GAGATGTACCTGGCGCTCATGAAGCATCACCTGGAGGGCATTTCGAAGCCGGGCGGACTGTTCGTCGACACCGCCGTCAG
CAAGCTGCCCTGGCGAGGACAACAGCGCCGCGTGCGGATGGTCGTCTACCGCCGGATCCGCAAGGAGGATGCGCAGATTC
GCGGACAGGACCCGGCGGCGTACCTGAAATCCATCTGCGAGCGTATCCAAGGCGGCCTGGCGAACGCCGGCATCGTCGCT
TCGCGCATGGGCGGACAGGAGATCAGGAACTGGTTGATCCGCTGGTTCAACCCGCACCCGGATCACCTCGGCCAGGCCGA
GGCGGACCTACGTCGCTTCTACGAACTGGTATGCCGTCCGGACGAACCGATCCTGCAGGATGAATTGCCACTGGCCGACG
GCACTGACTTCTCCCAGAACCTGTTCTATCGGCAGCCTGTTTCCGATGCCACCCAGGGCGTATGGCTCTTCGATGCCATG
CCGCACCGAGTGATTGTGGTCGACCAGTTGAACAAAGCGCCGCTGACAGGCCACTTCACCGGCGAGACGCTCAAAGGCGA
TGGCCTCAACGCCCTGTTCGATCGAATGCCCGAGGACACGCTGCTGTGCATCACCATGGTCGTGACGCCGCAGGACATGC
TGGAAGGGCATCTGCAGCAGCTCTCGAAAAAGGCCGTTGGTGACACCCAGGCCTCGATCCACACCCGCGAGGACGTGGCC
ACCGTTCGACGCCTGATCGGCCGGGAGCACAAGCTCTATCGCGGAGCGATCGCTCTGTTCGTGCGCGGCCGCGACCATAC
CCAGTTGGAGGAACGCTGCATCACCCTGAGCAACGTACTGCTCGGCGCCGGCCTGGTGCCGGTCGAACCGCAGAACGAAG
TCGGACCGCTGAACAGCTACCTGCGCTGGCTCCCCTCAAACTTCGATCCAAACGAGAAGCGAGCCCTGGAGTGGTACACC
CAGATGATGTTCGCTCAGCACATCGCCAACCTGTCGCCCATCTGGGGGCGCACCACCGGTACCGGACACCCTGGCTTCAC
GCTGTTCAACCGTGGCGGCGCGCCGTTGACCTTCGACCCGTTCAACAAGCTGGACCGGCAGATGAATGCCCACGGCTTCA
TCTTCGGGCCAACTGGCTCCGGCAAGTCGGCGTCCCTGACCAACCTCATCTGCCAGATGCTCGCCATGTACCTGCCGCGG
ATGTTCGTCGCGGAAGCGGGCAACAGCTTCGGCCTGCTGGCCGACTTAGCCAAGCGGTTTGGCCTCTCGGTCCACCGGGT
GCGCCTCGCCCCGGGCTCCGGCGTCAGCCTGGCGCCGTTCGCGGACGCCATCAAGCTGGTCGAGAGCCCCGACCAAGTGA
AGGTGCTGGACGCCGAAGACATCGAGGCCTCGGACTCGGTCCAGGGCAGCAAGGCCGACCTCGAGGACGACCAGCGAGAC
ATCCTGGGCGAGATGGAGATCGTCGCCCGCCTCATGATTACCGGTGGCGAAGAGAAGGAAGATGCGCGCCTGACCCGTGC
CGATCGCAGCGCCGTCCGCCAGGCGATCCTGGCGGCGGCCAGGACCTGCGCCGCCGCGAACCGCACGGTACTGACCCAAG
ACGTGCGCGATGCGCTCTACGAGGCCTCCAGGAGCGATAGCACCGCGCCAGAACGCCGCGCGCGGATCGCCGAAATGGCG
GAAGCCATGCAGATGTTCTGCATGGGCGCCGACGGCGAGATGTTCAATCGCGAAGGCACGCCCTGGCCTGAGGCCGACCT
TACCGTGGTGGATTTCGCAACGTACGCGCGCGAAGGCTACGCCGCCCAGCTCGGGATCGCCTACATCTCGCTGCTGAACA
CCGTGAACAACATCGCCGAACGCGACCAGTTCAAGGGCCGGCCAATCGTCAAGATCACCGATGAGGGGCACATCATCACC
AAGCACCCGCTGCTGCTGCCCTACGCCATGAAGATCACCAAGATGTGGCGGAAACTGGGCGCCTGGTTCTGGCTCGCCAC
CCAGAACATCGACGACATCCCAGCCTCCGGGGCGCCGATGCTGAACATGATCGAGTGGTGGTTGTGCCTGAACATGCCCC
CCGACGAAGTAGAGAAGATATCCAGGTTCCGCGAGCTGTCGCCGGCGCAGAAGTCGATGATGCTCTCGGCCCGCAAGGAA
AGCGGCAAGTTCACCGAGGGCGTGCTCCTGGCCAAGGGCAAAGAATACCTCGTCCGTGTGGTTCCCCCGAGTCTCTACCT
GGCCCTGGCCATGACCGAAAACGAAGAAAAGAACCAGCGCTACAACATCATGCAAGCCACCGGCTGCGACGAGCTCGAGG
CGGCCTTGCAGGTCGCAGCGGATCTCGACAAGGCGCGCGGCCTGCCACCCTTCCCCATTGTTTTCCCAGACCAACCGGCA
GTGGAGTGCCAGGACGAATGA

>ORF1040 (SEQ ID NO:14)
GTACCTGCCCGACGAGCAAGTGATGCTCCTGGAGGATGGGCGTTCGCGCGCCGCATTCTTCGAACTGGTGCCCTTGGGCA
CCGAGGGCCGCGATCCCAATTGGATGCAGAACGCCCGGGACGCATTGAAAGAAGCCCTGCAGAACTCCTTCGACGAGCAC
GAAACCTCACCCTGGATTGTCCAGTTCTACGCCCAGGACGAGATCAGCTGGGACAATTTCCAGGAGCAGTTGAGGCAGTA
CGTCCATCCTCGAGCGCGAGGATCGGCCTTCAGCGAGATGTACCTGGCGCTCATGAAGCATCACCTGGAGGGCATTTCGA
AGCCGGGCGGACTGTTCGTCGACACCGCCGTCAGCAAGCTGCCCTGGCGAGGACAACAGCGCCGCGTGCGGATGGTCGTC
TACCGCCGGATCCGCAAGGAGGATGCGCAGATTCGCGGACAGGACCCGGCGGCGTACCTGAAATCCATCTGCGAGCGTAT
CCAAGGCGGCCTGGCGAACGCCGGCATCGTCGCTTCGCGCATGGGCGGACAGGAGATCAGGAACTGGTTGATCCGCTGGT
TCAACCCGCACCCGGATCACCTCGGCCAGGCCGAGGCGGACCTACGTCGCTTCTACGAACTGGTATGCCGTCCGGACGAA
CCGATCCTGCAGGATGA

Fig. 3-2

>ORF1640c (SEQ ID NO:16)
GTCCGCCTCGGCCTGGCCGAGGTGATCCGGGTGCGGGTTGAACCAGCGGATCAACCAGTTCCTGATCTCCTGTCCGCCCA
TGCGCGAAGCGACGATGCCGGCGTTCGCCAGGCCGCCTTGGATACGCTCGCAGATGGATTTCAGGTACGCCGCCGGGTCC
TGTCCGCGAATCTGCGCATCCTCCTTGCGGATCCGGCGGTAGACGACCATCCGCACGCGGCGCTGTTGTCCTCGCCAGGG
CAGCTTGCTGACGGCGGTGTCGACGAACAGTCCGCCCGGCTTCGAAATGCCCTCCAGGTGATGCTTCATGAGCGCCAGGT
ACATCTCGCTGAAGGCCGATCCTCGCGCTCGAGGATGGACGTACTGCCTCAACTGCTCCTGGAAATTGTCCCAGCTGATC
TCGTCCTGGGCGTAGAACTGGACAATCCAGGGTGA

>ORF2228c (SEQ ID NO:18)
GGGGAGCCAGCGCAGGTAGCTGTTCAGCGGTCCGACTTCGTTCTGCGGTTCGACCGGCACCAGGCCGGCGCCGAGCAGTA
CGTTGCTCAGGGTGATGCAGCGTTCCTCCAACTGGGTATGGTCGCGGCCGCGCACGAACAGAGCGATCGCTCCGCGATAG
AGCTTGTGCTCCCGGCCGATCAGGCGTCGAACGGTGGCCACGTCCTCGCGGGTGTGGATCGAGGCCTGGGTGTCACCAAC
GGCCTTTTTCGAGAGCTGCTGCAGATGCCCTTCCAGCATGTCCTGCGGCGTCACGACCATGGTGATGCACAGCAGCGTGT
CCTCGGGCATTCGATCGAACAGGGCGTTGAGGCCATCGCCTTTGAGCGTCTCGCCGGTGAAGTGGCCTGTCAGCGGCGCT
TTGTTCAACTGGTCGACCACAATCACTCGGTGCGGCATGGCATCGAAGAGCCATACGCCCTGGGTGGCATCGGAAACAGG
CTGCCGATAGAACAGGTTCTGGGAGAAGTCAGTGCCGTCGGCCAGTGGCAATTCATCCTGCAGGATCGGTTCGTCCGGAC
GGCATACCAGTTCGTAGAAGCGACGTAG

>ORF2068c (SEQ ID NO:20)
AGCTTGTGCTCCCGGCCGATCAGGCGTCGAACGGTGGCCACGTCCTCGCGGGTGTGGATCGAGGCCTGGGTGTCACCAAC
GGCCTTTTTCGAGAGCTGCTGCAGATGCCCTTCCAGCATGTCCTGCGGCGTCACGACCATGGTGATGCACAGCAGCGTGT
CCTCGGGCATTCGATCGAACAGGGCGTTGAGGCCATCGCCTTTGAGCGTCTCGCCGGTGAAGTGGCCTGTCAGCGGCGCT
TTGTTCAACTGGTCGACCACAATCACTCGGTGCGGCATGGCATCGAAGAGCCATACGCCCTGGGTGGCATCGGAAACAGG
CTGCCGATAG

>ORF1997 (SEQ ID NO:22)
CACCCAGGCCTCGATCCACACCCGCGAGGACGTGGCCACCGTTCGACGCCTGATCGGCCGGGAGCACAAGCTCTATCGCG
GAGCGATCGCTCTGTTCGTGCGCGGCCGCGACCATACCCAGTTGGAGGAACGCTGCATCACCCTGAGCAACGTACTGCTC
GGCGCCGGCCTGGTGCCGGTCGAACCGCAGAACGAAGTCGGACCGCTGAACAGCTACCTGCGCTGGCTCCCCTCAAACTT
CGATCCAAACGAGAAGCGAGCCCTGGAGTGGTACACCCAGATGATGTTCGCTCAGCACATCGCCAACCTGTCGCCCATCT
GGGGGCGCACCACCGGTACCGGACACCCTGGCTTCACGCTGTTCAACCGTGGCGGCGCGCCGTTGACCTTCGACCCGTTC
AACAAGCTGGACCGGCAGATGAATGCCCACGGCTTCATCTTCGGGCCAACTGGCTCCGGCAAGTCGGCGTCCCTGACCAA
CCTCATCTGCCAGATGCTCGCCATGTACCTGCCGCGGATGTTCGTCGCGGAAGCGGGCAACAGCTTCGGCCTGCTGGCCG
ACTTAGCCAAGCGGTTTGGCCTCTCGGTCCACCGGGTGCGCCTCGCCCCGGGCTCCGGCGTCAGCCTGGCGCCGTTCGCG
GACGCCATCAAGCTGGTCGAGAGCCCCGACCAAGTGAAGGTGCTGGACGCCGAAGACATCGAGGCCTCGGACTCGGTCCA
GGGCAGCAAGGCCGACCTCGAGGACGACCAGCGAGACATCCTGGGCGAGATGGAGATCGTCGCCCGCCTCATGATTACCG
GTGGCGAAGAGAAGGAAGATGCGCGCCTGACCCGTGCCGATCGCAGCGCCGTCCGCCAGGCGATCCTGGCGGCGGCCAGG
ACCTGCGCCGCCGCGAACCGCACGGTACTGACCCAAGACGTGCGCGATGCGCTCTACGAGGCCTCCAGGAGCGATAG

>ORF2558c (SEQ ID NO:24)
GTCGGCCAGCAGGCCGAAGCTGTTGCCCGCTTCCGCGACGAACATCCGCGGCAGGTACATGGCGAGCATCTGGCAGATGA
GGTTGGTCAGGGACGCCGACTTGCCGGAGCCAGTTGGCCCGAAGATGAAGCCGTGGGCATTCATCTGCCGGTCCAGCTTG
TTGAACGGGTCGAAGGTCAACGGCGCGCCGCCACGGTTGAACAGCGTGAAGCCAGGGTGTCCGGTACCGGTGGTGCGCCC
CCAGATGGGCGACAGGTTGGCGATGTGCTGAGCGAACATCATCTGGGTGTACCACTCCAGGGCTCGCTTCTCGTTTGGAT
CGAAGTTTGA

Fig. 3-3

>ORF2929c (SEQ ID NO:26)
AGCGCATCGCGCACGTCTTGGGTCAGTACCGTGCGGTTCGCGGCGGCGCAGGTCCTGGCCGCCGCCAGGATCGCCTGGCG
GACGGCGCTGCGATCGGCACGGGTCAGGCGCGCATCTTCCTTCTCTTCGCCACCGGTAATCATGAGGCGGGCGACGATCT
CCATCTCGCCCAGGATGTCTCGCTGGTCGTCCTCGAGGTCGGCCTTGCTGCCCTGGACCGAGTCCGAGGCCTCGATGTCT
TCGGCGTCCAGCACCTTCACTTGGTCGGGGCTCTCGACCAGCTTGATGGCGTCCGCGAACGGCGCCAGGCTGACGCCGGA
GCCCGGGGCGAGGCGCACCCGGTGGACCGAGAGGCCAAACCGCTTGGCTAAGTCGGCCAGCAGGCCGAAGCTGTTGCCCG
CTTCCGCGACGAACATCCGCGGCAGGTACATGGCGAGCATCTGGCAGATGAGGTTGGTCAGGGACGCCGACTTGCCGGAG
CCAGTTGGCCCGAAGATGAAGCCGTGGGCATTCATCTGCCGGTCCAGCTTGTTGAACGGGTCGAAGGTCAACGGCGCGCC
GCCACGGTTGAACAGCGTGAAGCCAGGGTGTCCGGTACCGGTGGTGCGCCCCCAGATGGGCGACAGGTTGGCGATGTGCT
GA

>ORF3965c (SEQ ID NO:28)
GCGCCTGTTGGGCCGTATCAGGCTGTGGATGTTGTTGCAGCCATTCATCCAAGAGCTGCTTTATCTGCGGGACGATATCC
CGGCGATCGACTGCCCTCAGTTGAATCTGCTGCAGCTCCTCTATCAGTACAGGAGCGCATATCCTTAGCGTCTGCAGGGC
ATCCTCTTCGGGGTTCTGCAGGATCTGGGTCAGGTTGTCGATCAGGTTCTGGGTCAGCGAATTCAGAACTCTCATTCGTC
CTGGCACTCCACTGCCGGTTGGTCTGGGAAAACAATGGGGAAGGGTGGCAGGCCGCGCGCCTTGTCGAGATCCGCTGCGA
CCTGCAAGGCCGCCTCGAGCTCGTCGCAGCCGGTGGCTTGCATGATGTTGTAGCGCTGGTTCTTTTCTTCGTTTTCGGTC
ATGGCCAGGGCCAGGTAGAGACTCGGGGGAACCACACGGACGAGGTATTCTTTGCCCTTGGCCAGGAGCACGCCCTCGGT
GAACTTGCCGCTTTCCTTGCGGGCCGAGAGCATCATCGACTTCTGCGCCGGCGACAGCTCGCGGAACCTGGATATCTTCT
CTACTTCGTCGGGGGGCATGTTCAGGCACAACCACCACTCGATCATGTTCAGCATCGGCGCCCCGGAGGCTGGGATGTCG
TCGATGTTCTGGGTGGCGAGCCAGAACCAGGCGCCCAGTTTCCGCCACATCTTGGTGATCTTCATGGCGTAGGGCAGCAG
CAGCGGGTGCTTGGTGATGATGTGCCCCTCATCGGTGATCTTGACGATTGGCCGGCCCTTGAACTGGTCGCGTTCGGCGA
TGTTGTTCACGGTGTTCAGCAGCGAGATGTAGGCGATCCCGAGCTGGGCGGCGTAGCCTTCGCGCGCGTACGTTGCGAAA
TCCACCACGGTAAGGTCGGCCTCAGGCCAGGGCGTGCCTTCGCGATTGAACATCTCGCCGTCGGCGCCCATGCAGAACAT
CTGCATGGCTTCCGCCATTTCGGCGATCCGCGCGCGGCGTTCTGGCGCGGTGCTATCGCTCCTGGAGGCCTCGTAGAGCG
CATCGCGCACGTCTTGGGTCAGTACCGTGCGGTTCGCGGCGGCGCAGGTCCTGGCCGCCGCCAGGATCGCCTGGCGGACG
GCGCTGCGATCGGCACGGGTCAGGCGCGCATCTTCCTTCTCTTCGCCACCGGTAATCATGAGGCGGGCGACGATCTCCAT
CTCGCCCAGGATGTCTCGCTGGTCGTCCTCGAGGTCGGCCTTGCTGCCCTGGACCGAGTCCGAGGCCTCGATGTCTTCGG
CGTCCAGCACCTTCACTTGGTCGGGGCTCTCGACCAGCTTGATGGCGTCCGCGAACGGCGCCAGGCTGACGCCGGAGCCC
GGGGCGAGGCGCACCCGGTGGACCGAGAGGCCAAACCGCTTGGCTAA

>ORF3218 (SEQ ID NO:30)
GGGGCACATCATCACCAAGCACCCGCTGCTGCTGCCCTACGCCATGAAGATCACCAAGATGTGGCGGAAACTGGGCGCCT
GGTTCTGGCTCGCCACCCAGAACATCGACGACATCCCAGCCTCCGGGGCGCCGATGCTGAACATGATCGAGTGGTGGTTG
TGCCTGAACATGCCCCCCGACGAAGTAGAGAAGATATCCAGGTTCCGCGAGCTGTCGCCGGCGCAGAAGTCGATGATGCT
CTCGGCCCGCAAGGAAAGCGGCAAGTTCACCGAGGGCGTGCTCCTGGCCAAGGGCAAAGAATACCTCGTCCGTGTGGTTC
CCCCGAGTCTCTACCTGGCCCTGGCCATGACCGAAAACGAAGAAAAGAACCAGCGCTACAACATCATGCAAGCCACCGGC
TGCGACGAGCTCGAGGCGGCCTTGCAGGTCGCAGCGGATCTCGACAAGGCGCGCGGCCTGCCACCCTTCCCCATTGTTTT
CCCAGACCAACCGGCAGTGGAGTGCCAGGACGAATGAGAGTTCTGAATTCGCTGACCCAGAACCTGATCGACAACCTGAC
CCAGATCCTGCAGAACCCCGAAGAGGATGCCCTGCAGACGCTAAGGATATGCGCTCCTGTACTGATAGAGGAGCTGCAGC
AGATTCAACTGAGGGCAGTCGATCGCCGGGATATCGTCCCGCAGATAAAGCAGCTCTTGGATGAATGGCTGCAACAACAT
CCACAGCCTGATACGGCCCAACAGGCGCTCATTGAGGCCGTGGACCGCGCGGAGATCCTACAGCGGAGGCAAGCGTGA

>ORF3568 (SEQ ID NO:32)
CCGAAAACGAAGAAAAGAACCAGCGCTACAACATCATGCAAGCCACCGGCTGCGACGAGCTCGAGGCGGCCTTGCAGGTC
GCAGCGGATCTCGACAAGGCGCGCGGCCTGCCACCCTTCCCCATTGTTTTCCCAGACCAACCGGCAGTGGAGTGCCAGGA
CGAATGAGAGTTCTGAATTCGCTGACCCAGAACCTGATCGACAACCTGACCCAGATCCTGCAGAACCCCGAAGAGGATGC
CCTGCAGACGCTAAGGATATGCGCTCCTGTACTGATAGAGGAGCTGCAGCAGATTCAACTGAGGGCAGTCGATCGCCGGG
ATATCGTCCCGCAGATAAAGCAGCTCTTGGATGA

Fig. 3-4

>ORF4506c (SEQ ID NO:34)
GTCAATAAGTTCGTTGTCTTTCGCACATTTCTCCAGTCGAGCCTGGTCCAGTTCAGGAAAGTCCAATGTGCCGCCAGGCA
GCCCGCCCCCGTTGCCGGCCGACTGAGCGAAGATCGCATCGATAGCGCTCCAGAAGGCTTTGGCGCCGCCTTGGATCCCC
GCGCACTCCACCAGGCGAGCCTGGTGGCGGGCCGCCTCGCCATGCATCTGCAGGGGAAGATGGCGCCAAACCAGGTTCAC
GTCCGGATGGCTGTCTACCCAGCGCTTAAGCCGCGGGGTGTAGACCTTGCAGAAGGGGCACTCCAGGTCGGCGTATTCAT
TGATCGTCCAGCGCGCTTTCGCATCGCCGTAGAGGCTGTGGTTGGCTGGCAGGCCCTTCACCAGAAGCTCTACCCCTACG
GCGGATGCAGCCAGCAAGACCAGCAGCAGCCCCGCCCAGGGCAGGGCGGGACCTTGAAATCGTTTGGCTGCCCAGCCGCC
CTTCAAGAGTCTCACGCTTGCCTCCGCTGTAGGATCTCCGCGCGGTCCACGGCCTCAATGAGCGCCTGTTGGGCCGTATC
AGGCTGTGGATGTTGTTGCAGCCATTCATCCAAGAGCTGCTTTATCTGCGGGACGATATCCCGGCGATCGACTGCCCTCA
GTTGA

>ORF3973 (SEQ ID NO:36)
GGCCGTGGACCGCGCGGAGATCCTACAGCGGAGGCAAGCGTGAGACTCTTGAAGGGCGGCTGGGCAGCCAAACGATTTCA
AGGTCCCGCCCTGCCCTGGGCGGGGCTGCTGCTGGTCTTGCTGGCTGCATCCGCCGTAGGGGTAGAGCTTCTGGTGAAGG
GCCTGCCAGCCAACCACAGCCTCTACGGCGATGCGAAAGCGCGCTGGACGATCAATGAATACGCCGACCTGGAGTGCCCC
TTCTGCAAGGTCTACACCCCGCGGCTTAAGCGCTGGGTAGACAGCCATCCGGACGTGAACCTGGTTTGGCGCCATCTTCC
CCTGCAGATGCATGGCGAGGCGGCCCGCCACCAGGCTCGCCTGGTGGAGTGCGCGGGGATCCAAGGCGGCGCCAAAGCCT
TCTGGAGCGCTATCGATGCGATCTTCGCTCAGTCGGCCGGCAACGGGGGCGGGCTGCCTGGCGGCACATTGGACTTTCCT
GAACTGGACCAGGCTCGACTGGAGAAATGTGCGAAAGACAACGAACTTATTGACTCAGATATCAAGTTGGACATCGACAT
TGCACGGTCGAAGGGCATTACAGCGACCCCGACCCTCGTCATCCGGGACAACCAGACGGGACGAAGCGTGAAGCTTGAAG
GCATGGCCGACGAGACCACGTTGCTGTCGGCGATAGACTGGCTAGCCAAGGATCTCTAG

>ORF4271 (SEQ ID NO:38)
ACCTGGTTTGGCGCCATCTTCCCCTGCAGATGCATGGCGAGGCGGCCCGCCACCAGGCTCGCCTGGTGGAGTGCGCGGGG
ATCCAAGGCGGCGCCAAAGCCTTCTGGAGCGCTATCGATGCGATCTTCGCTCAGTCGGCCGGCAACGGGGGCGGGCTGCC
TGGCGGCACATTGGACTTTCCTGAACTGGACCAGGCTCGACTGGAGAAATGTGCGAAAGACAACGAACTTATTGACTCAG
ATATCAAGTTGGACATCGACATTGCACGGTCGAAGGGCATTACAGCGACCCCGACCCTCGTCATCCGGGACAACCAGACG
GGACGAAGCGTGA

>ORF4698 (SEQ ID NO:40)
GAAATCGGCGAGGATTCCAACATCCCTCTTTTGGTCCTCCAGGATGCCCTGCACTTCACCTGGCAGAACCTCGACCTCCT
CCCCATCCACAATCTTTACCATTCTCTTGTGGCCGGAGCTGGTGAGGCTAAGCCTCAACTCCATTGCCGGCCGAGCATTG
ATGTAAATGCTCTCGAGCAAGCGCTCCATGACTTCGACCACTCCTTAATATCAGTTAGCCAGCTACATACAGGAATTATG
CTACCCAGGACATGCAGGCGTCACCCCTACTTATGTACGTGGCAGCGTTCGATCACGGCTCGAAAAAATACACCACCTAC
GAGTTGA

>ORF5028 (SEQ ID NO:42)
TTTCCTGCTGCCCTATCGGAAGTGATCCTGTCTGCTGTCTGTACCTTTCTAGAACCGGTACAGACCCATGCCTCTTCATC
ACTCCCCCCCTGGCCGGCGGCCACCAACGCTGGCCGTTGGCGTACTACTGGTACTGCTGAGCAGCGCGAGTCAGGCCGAA
ACCTGGGTCATCACCGACAAGGCTCATCCGGTCTCTGCCACCGGATCGTCGCGCGTTCTGTTTCTGGACGCCCAGGAACA
CCTCGAGGAGCAACTGACTGCGGCCTTGCCCCAGGATCCACAGCATGCTCAAGCGGCGTTTAA

>ORF5080 (SEQ ID NO:44)
AACCGGTACAGACCCATGCCTCTTCATCACTCCCCCCCTGGCCGGCGGCCACCAACGCTGGCCGTTGGCGTACTACTGGT
ACTGCTGAGCAGCGCGAGTCAGGCCGAAACCTGGGTCATCACCGACAAGGCTCATCCGGTCTCTGCCACCGGATCGTCGC
GCGTTCTGTTTCTGGACGCCCAGGAACACCTCGAGGAGCAACTGACTGCGGCCTTGCCCCAGGATCCACAGCATGCTCAA
GCGGCGTTTAAGCGATTGCTACAAAGCCCCGATGGGCGCCGCCTGCAGGCAGAGCTGGTCAAGGCACAACAAGACGTCGC
CGATGCGTGGAGTCTCGGTGTCGAGAAGATCCCTGCCGTAGTAGTCGATAGGCAGTACGTGGTCTACGGCGAACCGGATG
TTTCGCGCGCTCTTGAGCTAATCGCCAAGGCCAGGAGGTCGCGCTGA

Fig. 3-5

>ORF6479c (SEQ ID NO:46)
TTCGTCTCCGTGTCCTTATTGGAAGTCGGTACTGCAGATGAACATCTGCCCCTTGCGCTGGCAGCAGGAGTAGGGACGCC
AGAGCGCCCAGGCGTGCTCCCCGTCGACGGCTTGCGTCTTCGGCCCAGAGTTGGGAAACACCGCGCAGTTGAGGCTCAGG
GATGGGGTCAGCTCCTGCCATTTCCCGGTCGAGGCATCGCCCTCTTTCAGCTCGCCCGCCGGCCAGTAGCCGTCCTTGGG
GGCTGCGCGCATGGGGAGGTAGACGTGGAGCTGGCCGATTCGCGTGGTGATATCGCCGGCGCGCTGGGCGATGACGGCTG
CCGTCTTGTAGTCGTCGGTCTGGTGCAGGAAGCCGCTGCGCGGATAGAGGTTCCCCCACATGTCGCCGGAGAAGATTCCA
CCCACCTCGCGCAGCCCTGGGACCAACGCTTCGGGGTACACCTGCTCGGGAATTCCATGCCGCCAGCCAATGGCGTCCAG
TGTGCTGAGAAAGTACGGCACCAGCGGGACGGTGGCGCCAGGGCAAACGTACCCAGAGGCGCTGGCGAACCGGCTGAACG
TGGCGCCACCAGGATGGCCGATCACATCCGCTTCCTTGAAGCGGCCGATGCTGTTCTCGGCCTTGTAGTTTGTGGTCGCG
TCATTGCCGGCCTGGGCGAGTGGATTCGGTGTACCCAGCGCCGATACCTCGGTCCAGGGGTTGCTCCCGGTATTCGCGTA
GCTGGAGACGACTGCGTCAGGCACGTAGTGGCGGACCTTGACCGACGTCTTCACTTTGCAGCCATGCGGGCCGCAGAGCA
GCCAGTAACAGATCCCGACGACCTTGTATTCGAGGCACTGAGGGGAAAGGGTGGAGGAGACGATGGCAGCGCTGTTGATC
GCGGCCGAGGCCGTGAACGAGAGGCTGAAGGTGGCGGCCGCCGCTGCCAGGCGGCGGAGGTTGAGGCTGGTCATCAGCGC
GACCTCCTGGCCTTGGCGATTAGCTCAAGAGCGCGCGAAACATCCGGTTCGCCGTAG

>ORF5496 (SEQ ID NO:48)
GCTAATCGCCAAGGCCAGGAGGTCGCGCTGATGACCAGCCTCAACCTCCGCCGCCTGGCAGCGGCGGCCGCCACCTTCAG
CCTCTCGTTCACGGCCTCGGCCGCGATCAACAGCGCTGCCATCGTCTCCTCCACCCTTTCCCCTCAGTGCCTCGAATACA
AGGTCGTCGGGATCTGTTACTGGCTGCTCTGCGGCCCGCATGGCTGCAAAGTGAAGACGTCGGTCAAGGTCCGCCACTAC
GTGCCTGACGCAGTCGTCTCCAGCTACGCGAATACCGGGAGCAACCCCTGGACCGAGGTATCGGCGCTGGGTACACCGAA
TCCACTCGCCCAGGCCGGCAATGACGCGACCACAAACTACAAGGCCGAGAACAGCATCGGCCGCTTCAAGGAAGCGGATG
TGATCGGCCATCCTGGTGGCGCCACGTTCAGCCGGTTCGCCAGCGCCTCTGGGTACGTTTGCCCTGGCGCCACCGTCCCG
CTGGTGCCGTACTTTCTCAGCACACTGGACGCCATTGGCTGGCGGCATGGAATTCCCGAGCAGGTGTACCCCGAAGCGTT
GGTCCCAGGGCTGCGCGAGGTGGGTGGAATCTTCTCCGGCGACATGTGGGGGAACCTCTATCCGCGCAGCGGCTTCCTGC
ACCAGACCGACGACTACAAGACGGCAGCCGTCATCGCCCAGCGCGCCGGCGATATCACCACGCGAATCGGCCAGCTCCAC
GTCTACCTCCCCATGCGCGCAGCCCCCAAGGACGGCTACTGGCCGGCGGGCGAGCTGAAAGAGGGCGATGCCTCGACCGG
GAAATGGCAGGAGCTGACCCCCATCCCTGAGCCTCAACTGCGCGGTGTTTCCCAACTCTGGGCCGAAGACGCAAGCCGTCG
ACGGGGAGCACGCCTGGGCGCTCTGGCGTCCCTACTCCTGCTGCCAGCGCAAGGGGCAGATGTTCATCTGCAGTACCGAC
TTCCAATAA

Fig. 3-6

>ORF5840 (SEQ ID NO:50)
CGCGACCACAAACTACAAGGCCGAGAACAGCATCGGCCGCTTCAAGGAAGCGGATGTGATCGGCCATCCTGGTGGCGCCA
CGTTCAGCCGGTTCGCCAGCGCCTCTGGGTACGTTTGCCCTGGCGCCACCGTCCCGCTGGTGCCGTACTTTCTCAGCACA
CTGGACGCCATTGGCTGGCGGCATGGAATTCCCGAGCAGGTGTACCCCGAAGCGTTGGTCCCCAGGGCTGCGCGAGGTGGG
TGGAATCTTCTCCGGCGACATGTGGGGGAACCTCTATCCGCGCAGCGGCTTCCTGCACCAGACCGACGACTACAAGACGG
CAGCCGTCATCGCCCAGCGCGCCGGCGATATCACCACGCGAATCGGCCAGCTCCACGTCTACCTCCCCATGCGCGCAGCC
CCCAAGGACGGCTACTGGCCGGCGGGCGAGCTGAAAGAGGGCGATGCCTCGACCGGGAAATGGCAGGAGCTGACCCCATC
CCTGAGCCTCAACTGCGCGGTGTTTCCCAACTCTGGGCCGAAGACGCAAGCCGTCGACGGGGAGCACGCCTGGGCGCTCT
GGCGTCCCTACTCCTGCTGCCAGCGCAAGGGGCAGATGTTCATCTGCAGTACCGACTTCCAATAAGGACACGGAGACGAA
TCATGCGAATGAACATCACCTCGGTCGCGCTAATGTGGCTGCTCGCAGCGCAACTTGCCCAGGCCGACGACCCGATCAAC
GTGTCCAAGACCGGCACGGTGCTCAGCGACGAGGTCCTCTACAGCATTGGCGGCGGCAGTGCGGTGAGCATGGGCAGCGC
CGGCCAGATGGACTCGATCGGCGTCGGCTTCGGCTGGAACAACGACATGATGTGCGGAAACATGAACCTGAGCACCACCC
TGGAGAACCAGCTCAACGGTGCCACACAGGGTTTCCAGAACATCATGGGCTCAGTCATCCAGAACGCGACCGGCGCGGTC
ATGTCGCTGCCGGCGTTGATCATCCAGCGCGCGAACCCTCAGCTCTACAACCTGATCACCAATGGCATCCTGCAGGCGCG
GATCGACTACGACCGCTCGAAAGGGACTTGCAAAACGATCGCCGAAAAGATGGCTGACATCGCTGGCGAGCAGACCGGCT
GGGGGAAAAATCGCCGAAGGCCAAGCCCTGGGCGCCACACTGGCCTCTGACGGGAAAGACGCCGTATCCGCCCTCGAAGCA
GTGGAGAAGAAAGGCGGCAACGATGGCGTAACCTGGGTTGGTGGAGACAAGGCCGGCGGCTCCGGCCAGAAGCCCATTCG
CATCGTCAACGACGTGACCCGGGCGGGCTACAACCTGTTGACCAGCCGCTCGGTGAATGATTCGTCGAGCGTGCCTTCCG
CCACTTGCAACAACGGCCTGGTCTGCAACACTTGGTCCTCCCCCCAGGAGGCCGCCGCATTCGCCACCCGGGTACTGGGG
GAGCAACAGCAACAGACCTGCGAAGGCTGCCAGAAGACGGTGACGGCTGCTGGCGTCGGCCTCACCCCGCTGATCCAGGA
GACCTACGACAAGAAGCTCCAGTCGCTGCAGGAGCTGCTGTCGAAGAGCAAACCACTGACTGCAGAGAACCTGGCTGCGG
CCGGCACCGATGCTCTGCCAATTACCCGCGGCGTCATCGAGGCGCTGCGCGACGAGCGTGACCAGGACGTCCTGGCGCGC
CGCCTGGCGTCCGATGTCTCCCTGATGGACGTGCTCAGCAAGGCACTGCTACTGCAGCGCCTGATGTTCGCCGGCGCCAA
GGAGCCCAACGTCGCCGCCAACGGCCTGGCCACCCAAGCCGTCGATCAGCAGACCAGCCTCCTGCAGCAGGAGATCTCCA
ATCTCAAGACCGAACTGGAACTCCGTCGCGAGTTGGCCAGCAACTCCCCCATGCGGGTCATCGAGCGCGGGCAACAACGC
GCCTCAGGGTCCAGTGGCGTGTTCGAGTCGGCGCCCGATGCCGATCGCCTCGATCGCCTGCAGGCCCCCTCTGCCGCCGG
CGGCAAGTCGGGAGGGAGACCGTGA

>ORF5899 (SEQ ID NO:52)
TCGGCCATCCTGGTGGCGCCACGTTCAGCCGGTTCGCCAGCGCCTCTGGGTACGTTTGCCCTGGCGCCACCGTCCCGCTG
GTGCCGTACTTTCTCAGCACACTGGACGCCATTGGCTGGCGGCATGGAATTCCCGAGCAGGTGTACCCCGAAGCGTTGGT
CCCAGGGCTGCGCGAGGTGGGTGGAATCTTCTCCGGCGACATGTGGGGGAACCTCTATCCGCGCAGCGGCTTCCTGCACC
AGACCGACGACTACAAGACGGCAGCCGTCATCGCCCAGCGCGCCGGCGATATCACCACGCGAATCGGCCAGCTCCACGTC
TACCTCCCCATGCGCGCAGCCCCCAAGGACGGCTACTGGCCGGCGGGCGAGCTGA

>ORF6325 (SEQ ID NO:54)
GCCTCAACTGCGCGGTGTTTCCCAACTCTGGGCCGAAGACGCAAGCCGTCGACGGGGAGCACGCCTGGGCGCTCTGGCGT
CCCTACTCCTGCTGCCAGCGCAAGGGGCAGATGTTCATCTGCAGTACCGACTTCCAATAAGGACACGGAGACGAATCATG
CGAATGAACATCACCTCGGTCGCGCTAATGTGGCTGCTCGCAGCGCAACTTGCCCAGGCCGACGACCCGATCAACGTGTC
CAAGACCGGCACGGTGCTCAGCGACGAGGTCCTCTACAGCATTGGCGGCGGCAGTGCGGTGAGCATGGGCAGCGCCGGCC
AGATGGACTCGATCGGCGTCGGCTTCGGCTGGAACAACGACATGATGTGCGGAAACATGAACCTGAGCACCACCCTGGAG
AACCAGCTCAACGGTGCCACACAGGGTTTCCAGAACATCATGGGCTCAGTCATCCAGAACGCGACCGGCGCGGTCATGTC
GCTGCCGGCGTTGATCATCCAGCGCGCGAACCCTCAGCTCTACAACCTGATCACCAATGGCATCCTGCAGGCGCGGATCG
ACTACGACCGCTCGAAAGGGACTTGCAAAACGATCGCCGAAAAGATGGCTGA

Fig. 3-7

>ORF7567c (SEQ ID NO:56)
CAGTGCCTTGCTGAGCACGTCCATCAGGGAGACATCGGACGCCAGGCGGCGCGCCAGGACGTCCTGGTCACGCTCGTCGC
GCAGCGCCTCGATGACGCCGCGGGTAATTGGCAGAGCATCGGTGCCGGCCGCAGCCAGGTTCTCTGCAGTCAGTGGTTTG
CTCTTCGACAGCAGCTCCTGCAGCGACTGGAGCTTCTTGTCGTAGGTCTCCTGGATCAGCGGGGTGAGGCCGACGCCAGC
AGCCGTCACCGTCTTCTGGCAGCCTTCGCAGGTCTGTTGCTGTTGCTCCCCCAGTACCCGGGTGGCGAATGCGGCGGCCT
CCTGGGGGGGAGGACCAAGTGTTGCAGACCAGGCCGTTGTTGCAAGTGGCGGAAGGCACGCTCGACGAATCATTCACCGAG
CGGCTGGTCAACAGGTTGTAGCCCGCCCGGGTCACGTCGTTGACGATGCGAATGGGCTTCTGGCCGGAGCCGCCGGCCTT
GTCTCCACCAACCCAGGTTACGCCATCGTTGCCGCCTTTCTTCTCCACTGCTTCGAGGGCGGATACGGCGTCTTTCCCGT
CAGAGGCCAGTGTGGCGCCCAGGGCTTGGCCTTCGGCGATTTTCCCCCAGCCGGTCTGCTCGCCAGCGATGTCAGCCATC
TTTTCGGCGATCGTTTTGCAAGTCCCTTTCGAGCGGTCGTAGTCGATCCGCGCCTGCAGGATGCCATTGGTGATCAGGTT
GTAGAGCTGAGGGTTCGCGCGCTGGATGATCAACGCCGGCAGCGACATGACCGCGCCGGTCGCGTTCTGGATGACTGA

>ORF7180 (SEQ ID NO:58)
TTCGTCGAGCGTGCCTTCCGCCACTTGCAACAACGGCCTGGTCTGCAACACTTGGTCCTCCCCCCAGGAGGCCGCCGCAT
TCGCCACCCGGGTACTGGGGGAGCAACAGCAACAGACCTGCGAAGGCTGCCAGAAGACGGTGACGGCTGCTGGCGTCGGC
CTCACCCCGCTGATCCAGGAGACCTACGACAAGAAGCTCCAGTCGCTGCAGGAGCTGCTGTCGAAGAGCAAACCACTGAC
TGCAGAGAACCTGGCTGCGGCCGGCACCGATGCTCTGCCAATTACCCGCGGCGTCATCGAGGCGCTGCGCGACGAGCGTG
A

>ORF7501 (SEQ ID NO:60)
CCAGGACGTCCTGGCGCGCCGCCTGGCGTCCGATGTCTCCCTGATGGACGTGCTCAGCAAGGCACTGCTACTGCAGCGCC
TGATGTTCGCCGGCGCCAAGGAGCCCAACGTCGCCGCCAACGGCCTGGCCACCCAAGCCGTCGATCAGCAGACCAGCCTC
CTGCAGCAGGAGATCTCCAATCTCAAGACCGAACTGGAACTCCGTCGCGAGTTGGCCAGCAACTCCCCCATGCGGGTCAT
CGAGCGCGGGCAACAACGCGCCTCAGGGTCCAGTGGCGTGTTCGAGTCGGCGCCCGATGCCGATCGCCTCGATCGCCTGC
AGGCCCCCTCTGCCGCCGGCGGCAAGTCGGGAGGGAGACCGTGATGGCAGATACGCTCACCACCCGAAAGCTTCTCGGTC
AGCTACTGGTCGGAGTGCTGATCGTCATCGGACTGGCAGTGGTCGGTACGCTGCTCAGTCTCTTCGCCCTGAACCACTTC
GGTGGCATCCAGGGCCTGGAGGCCTGGCGGCAAAGCAACTACTGGAGCTTGTTCGCCTGGCGGGCGCTGCTGTACTGCGC
CCTGGCCATCGCCTGGTTCCGGCAGCGCAAGGAACTGAGCGCGCATGAGCGGCAGCGCATTCGGCGGATCGAGATCCTGG
TGCTGTTGCTGGTCCTGCTCATCGAATTCAGCAAAGCCTACTTCCGCACGGGAGGCGCAGCATGA

>ORF7584 (SEQ ID NO:62)
TGTTCGCCGGCGCCAAGGAGCCCAACGTCGCCGCCAACGGCCTGGCCACCCAAGCCGTCGATCAGCAGACCAGCCTCCTG
CAGCAGGAGATCTCCAATCTCAAGACCGAACTGGAACTCCGTCGCGAGTTGGCCAGCAACTCCCCCATGCGGGTCATCGA
GCGCGGGCAACAACGCGCCTCAGGGTCCAGTGGCGTGTTCGAGTCGGCGCCCGATGCCGATCGCCTCGATCGCCTGCAGG
CCCCCTCTGCCGCCGGCGGCAAGTCGGGAGGGAGACCGTGATGGCAGATACGCTCACCACCCGAAAGCTTCTCGGTCAGC
TACTGGTCGGAGTGCTGATCGTCATCGGACTGGCAGTGGTCGGTACGCTGCTCAGTCTCTTCGCCCTGAACCACTTCGGT
GGCATCCAGGGCCTGGAGGCCTGGCGGCAAAGCAACTACTGGAGCTTGTTCGCCTGGCGGGCGCTGCTGTACTGCGCCCT
GGCCATCGCCTGGTTCCGGCAGCGCAAGGAACTGAGCGCGCATGA

>ORF8208c (SEQ ID NO:64)
AGGTCATGCTGCGCCTCCCGTGCGGAAGTAGGCTTTGCTGAATTCGATGAGCAGGACCAGCAACAGCACCAGGATCTCGA
TCCGCCGAATGCGCTGCCGCTCATGCGCGCTCAGTTCCTTGCGCTGCCGGAACCAGGCGATGGCCAGGGCGCAGTACAGC
AGCGCCCGCCAGGCGAACAAGCTCCAGTAGTTGCTTTGCCGCCAGGCCTCCAGGCCCTGGATGCCACCGAAGTGGTTCAG
GGCGAAGAGACTGAGCAGCGTACCGACCACTGCCAGTCCGATGACGATCAGCACTCCGACCAGTAG

Fig. 3-8

>ORF8109 (SEQ ID NO:66)
GCGGCAGCGCATTCGGCGGATCGAGATCCTGGTGCTGTTGCTGGTCCTGCTCATCGAATTCAGCAAAGCCTACTTCCGCA
CGGGAGGCGCAGCATGACCTTCATGACCAATGACTACCTGGAGTATTACCTCACCCTCCTCGGCTGGATCATCAACAACG
GGATCTGGAACATGATCTCGGATACTGGCCTGTTCGCGGTGCCGTTCGCGGCCATCGTGATGCGCGAATGGCTGAAAGTT
CGTGGGGAAGGCGCCGACGAGGGCAACAAGGGAGTGCTGTCTCTCGCCCGCATCGAGACGCATATCTACGTCGGCTACAT
CGTGGTCGCCCTGGCGGGGATCCCGGTCGTCAACGTGAGCTTCGACACCATCGAGTTCGACCAGACTCGCGCCCAGCAGT
GCCAATACAATCTGCCGGCACCGGCGGACACCGGCTGGTCGAGCTCCTTCAGCAGCCTGGCCGGCAAGAGTGCGCAGATG
CCGCTCTGGTGGGCGATGATGCACGCCCTGTCCAAGGGCTTCACCAGCGGCGCCATCGCGGCCATTCCGTGCGGCACGGA
TCTGCGGCAGATGCGAATGGAAGTGGACAACACGCGCGTGAACAATCCGCTGCTGGCACAAGAAATCGCTGATTTTTCCA
GAGACTGCTACGGGCCTTCCCGTGCGCGGCTGTTCATGCGGCAACCCGACCTGGGCTCCGTCGCCGAGGACAACAAGGCG
TTGCAAGACCTGAACTGGATCGGCTCCCGATTCTTGTTGAACACCCCGGGGTACTACGACACCGACTACTCGAAGAGTCC
CCGTCAGTCGTGGCCCTACAACGCCACCCGCGATGCCGGCCTGCCTCAGGTGGGCGGTGGTGGCGGCTACCCAACCTGCA
AGCAGTGGTGGGCTGACTCAGGGATCGGCTTGCGTGATCGGATCAAGGACCAGGTGGATCCGGACCTGATGACCAGCTTC
CTCAAGTGGGCGAAATGGTTGAACCAGGACGAGGTGACCGAGGCTGTCATTCGCCAGGTGATCTCACCCTCCAGCCAGGT
CAAGGGTAACGTCTACACCGATTACGGCGGGCAGGTGGGCGGCACCGTGTGGAACGGCATCGCGAGAACCGCAGGAACCT
TCGGCGTTGCGGTGGGCAGCTTGGCATACTTCCCGGCGATGGATATGGTCCGCCAGGCACTGCCGATGGTGATGTCGTTC
CTGAAGATGGCAATGGTCATCTGCATTCCGATGGTCCTGGTCATCGGCACCTATCAACTGAAAGTTGCCATGACGATGAC
GGTCGTCTTCTTTGCGATGATGTTCGTCGACTTCTGGTTTCAGTTAGCCAGATATATCGACAGCACGATACTTGATGCTT
TCTATGGTTCGGGATCACCACATCTTTCATTCAACCCAGTCATGGGGCTGAATACGGCTACTCAAGATGCGATCTTGAAC
TTCGTTATGGGTTCTATGTTCATTGTTTTACCACTACTGTGGATGACAGCGATCGGCTGGTCCGGAATTCAAGCAGGGTC
TGTTCTGAACGGATTGAGCAGAGGGACTGAAGGAGTTCAAGCCGCCGGCAAGGAAGCAGGAAATAGAGTTAAAAACGCAG
TTTGA

>ORF9005c (SEQ ID NO:68)
GTCAGCCCACCACTGCTTGCAGGTTGGGTAGCCGCCACCACCGCCCACCTGAGGCAGGCCGGCATCGCGGGTGGCGTTGT
AGGGCCACGACTGACGGGGACTCTTCGAGTAGTCGGTGTCGTAGTACCCCGGGGTGTTCAACAAGAATCGGGAGCCGATC
CAGTTCAGGTCTTGCAACGCCTTGTTGTCCTCGGCGACGGAGCCCAGGTCGGGTTGCCGCATGAACAGCCGCGCACGGGA
AGGCCCGTAGCAGTCTCTGGAAAAAATCAGCGATTTCTTGTGCCAGCAGCGGATTGTTCACGCGCGTGTTGTCCACTTCCA
TTCGCATCTGCCGCAGATCCGTGCCGCACGGAATGGCCGCGATGGCGCCGCTGGTGAAGCCCTTGGACAGGGCGTGCATC
ATCGCCCACCAGAGCGGCATCTGCGCACTCTTGCCGGCCAGGCTGCTGAAGGAGCTCGACCAGCCGGTGTCCGCCGGTGC
CGGCAGATTGTATTGGCACTGCTGGGCGCGAGTCTGGTCGAACTCGATGGTGTCGAAGCTCACGTTGACGACCGGGATCC
CCGCCAGGGCGACCACGATGTAGCCGACGTAGATATGCGTCTCGATGCGGGCGAGAGACAGCACTCCCTTGTTGCCCTCG
TCGGCGCCTTCCCCACGAACTTTCAGCCATTCGCGCATCACGATGGCCGCGAACGGCACCGCGAACAGGCCAGTATCCGA
GATCATGTTCCAGATCCCGTTGTTGATGATCCAGCCGAGGAGGGTGAGGTAATACTCCAGGTAGTCATTGGTCATGAAGG
TCATGCTGCGCCTCCCGTGCGGAAGTAG

>ORF8222 (SEQ ID NO:70)
CTACCTGGAGTATTACCTCACCCTCCTCGGCTGGATCATCAACAACGGGATCTGGAACATGATCTCGGATACTGGCCTGT
TCGCGGTGCCGTTCGCGGCCATCGTGATGCGCGAATGGCTGAAAGTTCGTGGGGAAGGCGCCGACGAGGGCAACAAGGGA
GTGCTGTCTCTCGCCCGCATCGAGACGCATATCTACGTCGGCTACATCGTGGTCGCCCTGGCGGGGATCCCGGTCGTCAA
CGTGAGCTTCGACACCATCGAGTTCGACCAGACTCGCGCCCAGCAGTGCCAATACAATCTGCCGGCACCGGCGGACACCG
GCTGGTCGAGCTCCTTCAGCAGCCTGGCCGGCAAGAGTGCGCAGATGCCGCTCTGGTGGGCGATGATGCACGCCCTGTCC
AAGGGCTTCACCAGCGGCGCCATCGCGGCCATTCCGTGCGGCACGGATCTGCGGCAGATGCGAATGGAAGTGGACAACAC
GCGCGTGAACAATCCGCTGCTGGCACAAGAAATCGCTGA

>ORF8755c (SEQ ID NO:72)
CAGTCTCTGGAAAAAATCAGCGATTTCTTGTGCCAGCAGCGGATTGTTCACGCGCGTGTTGTCCACTTCCATTCGCATCTG
CCGCAGATCCGTGCCGCACGGAATGGCCGCGATGGCGCCGCTGGTGAAGCCCTTGGACAGGGCGTGCATCATCGCCCACC
AGAGCGGCATCTGCGCACTCTTGCCGGCCAGGCTGCTGAAGGAGCTCGACCAGCCGGTGTCCGCCGGTGCCGGCAGATTG
TATTGGCACTGCTGGGCGCGAGTCTGGTCGAACTCGATGGTGTCGAAGCTCACGTTGACGACCGGGATCCCCGCCAGGGC
GACCACGATGTAG

Fig. 3-9

>ORF10475 (SEQ ID NO:86)
AGCATGGCCAAGTACCGCATCTCTCATGATGCCCAAGCGGACATCGTCGATATCCTGCGCTTCACCCACAACCACTTCGG
CGATGCCGCGCGCCGACGTTACCAGGCACTCATAGGGGCGGCGCTGGAAGCAGTTGCGACAGACCCACAACAGGTAGGCA
GCATCAGCCGTGAAGAACTGGGAGCTGGCCTGCGCAGCATCCACCTCGTTTACTGCCACTCGATGCCCAATGTCGGTAAG
GTTGTTCGGCCCAGGCACTTCGTCTTCTACCGGGTGGCGACAGACCAGGTGCTAGAGGTGGTTCGCGTGCTTCACGACGC
CATGGATGTGGATCAACACCTGCCCCAACGATGA

>ORF11095c (SEQ ID NO:88)
AGCCGCATGCAAGCGGTGGTCAGCACGAATGCAAATGCTTGGTCAGGGGGAATGCAATCGAGTGGTCAAGCCACTGCTAT
TGCGCATCAACCATGGGGCACCTGCTGGTGGATGTTCACCCGTAGCCTTTTCGTGTTCGCCGGCGCGAACGCAGCCCTTT
CTGCCTTCCGGCAGGCCCTTTCGGGTAGGGCTTTTACCCTTGTGAACCATTCCCTTCGCCCTTCAAGCCCATTTCCCCTT
TGGGCCATTTGCTCCTGTTACAGTTGCTCATCGTTGGGGCAGGTGTTGATCCACATCCATGGCGTCGTGAAGCACGCGAA
CCACCTCTAG

>ORF11264 (SEQ ID NO:90)
ACCGCGGTGCGGAGAGATCTCCTCAAACTGATGGGGTTGCACGCATATCGAAGCAGATTACATAGGAGGCTTGCGCTGTTC
AACAGCTCCTGAGGGGACTTGGGTTGCCCATGGTTTCCACGGCCCAATCGTTGACGTCATTGACGATTCCGCTGGCTTTT
TCAGTACGCATCGCTTGGCGCTCCATTACCCAGCCCAATGCGGCCTTGCCGTTGACCAAGCGATTCCAAGGACTGCGATC
CATGTAGCCAGCCCTCTAATGCATGTATGTATAGGTAAGGTCGTCGTTATTTCGGCGTGGATGTGCTGA

>ORF11738 (SEQ ID NO:92)
GAAGAGGTGATCATGAAGTTACAGGCATATCGGCTGCAGAACTACCGCCGGCTGCGCGATGTTGTCATCGAGCTCGATGA
CGAAATTTCTATCTTTGTCGGTGCCAACAACAGCGGGAAGACATCCGCCGTCCAAGGCCTGTACTCAATGCTTCGCGGCG
AAGTGAAGAAGTTCGAGCTCTTTGACTTCAGTGCGGCGCTGTGGGCCGAGATCGATGCGGTCGGCAGGACGCCCCCTGGC
GATGAGGATGCGCCCAAAAGGTTACCGTCCATACTCTTGGATCTCTGGTTCCGCGTCGGTGAAGACGACCTCGCCACTGC
GATGTCGCTGCTGCCGAGCACTGAGTGGGACGGCAAGTGCGTCGGGATCCGGGTAGCGTTCGAGCCTCGGGATGCCCACG
AGCTCGTCTGGAAGTTCCATGAACTACATGAGAAGGCCAACAACGCAGCTGTCGCGCTTGCGGCCAAGCGCAAGGCCGCC
GGGGAGCAAGCTGTGGAGGCGGGCGCGGAAGACGCGGCTGCGGTGGTGGCCGATGCCGGCGAGTACAAGCCTTGGCCAGA
AAGCCTGACGAAGTACCTCACAAAGGAACTGAGCAAGGAATACACCTTCCGCTACTACGTGCTCGATGAGCGGGCTTTTG
TCGGCTATCAGGCAAGGGAGGCCGACTACGAGCCGCTACCCCTAGGCAAGGAGCCGGGCGGTGCAGCCATTCTCAAGTCG
CTGGTGAGGGTCGACTTCCTGCGCGCGCAGCGGCACCTCGATGACCCAGATGCCGGTAGCTCTGATCGCGCAGAGAGCTT
GTCGCGGCGTCTGAGCAGGTTCTATCACCGCAACCTGGAGAAGCGTGGCGACGACCATGCGGCTCTCAAGGCGCTAGATA
CCTCGGAGAAGGAGCTGAACTTCCACCTGAAGGAAGTCTTCAATGACACCCTCACGCGCCTGGCCAAGCTCGGCTATCCG
GGCGTCAACAATCCGGAGATCGTGATTCGGGCGGCCTTGGATCCGACCACTGTCTTGGGGCAAGACGCCAAGGTTCACTA
CGTGATCCCGGGCGTAGCTTCCGCCCAACTGCCAGACAGCTACAATGGCCTGGGGTTCAAGAATCTGGTCTACATGGTGG
TTGAGCTGCTCGACTTGCACGAGCAGTGGAAAGCCGAGGATGACAAGCGAGCTCCGCTTCATTTGGTCTTCATTGAGGAG
CCTGAGGCGCATCTGCACGCGCAGATCCAGCAGGTCTTCATCAGGAACGTTTTGCGCCTCCTTGAGGATGCTAACGATCA
CGCGACTTTGTTCCACACGCAGCTCGTCATCACCACGCACTCCCCGCACATCCTCTATGAACGCGGATTCTCGCCCATTC
GGTACTTCCGCCGCGTCAACGACCAGTTGGGCCATCACACGGATGTGCGCAATCTGTCGCTATTCAAAACGGGCGCGTCC
GACGCTCCAGCGCGCGAATTCCTGCAGCGGTATCTGAAGCTGACGCACTGCGATCTCTTTTTTTCCGACGCGGTGATATT
GGTGGAAGGCAACGTCGAGCGTCTGCTCCTGCCTGCAATGATCGAGTTGGTGGCCAAGCGCCTGCGTTCTTCCGCCCTAA
CCATCCTTGAAGTCGGTGGTGCGTTCGCGCATCGGTTCCAGGAGCTGATCGCCTTCGTTGGGCTCACAACACTGGTCATC
ACGGATCTGGACAGCGTGACGGTCAAGACGGACGCCGAGAAGGCCGCCGCGCAAGGCGCAGGCGCTGAGGGCGCCGTTGA
CGGAGATGACGAGGACGAGGACGACGACCTGAAGCCCTTCGAGCTTGAAGACGACGACGAAGCAGAACCGAGTGGCAAGA
AGAAGTCCAAGAAGCGTGGCAGCACCTGCCATGCACACGTGGAAGGTGCCGTCACGTCCAACCAAACCCTCATCAGCTGG
ATCCCGAAGAAGCGGTCGATGGCAGAGCTCTGGGAAGTCACGGCGGAGCAAAAGACGCTGTCGCTGGCTGAGGATTCCAG
CGCTGGGGTTCGGGTAGCTTACCAGACCAAGGTTTCGGTGACGGTGGGTGCGACGACATCACAGCTCTGCGGCCGCACAC
TTGAGGAGGCCTTTGGTCTTGAGAACGCGGACTGGTGCCAGGCTGAGGCAAACCGGTCGGTCGGCCTCAAGCTCAAGCGC
GCACCGAGCAGCCCTGAAGAGCTGGCTGAGAAGTTACACGATAGGGTGGTCGGCAAGAACTTCGACAAGACCCGCTTTGC
GCTGGAGGTACTCGCAAGCGGGCCGCTCAATGGCTGGAAGGTTCCCGCGTACATCGCCGAGGGCTTGGCCTGGCTCGAAG
CCAAAGTGGCCCACGAGCTTGAGGCGGATGCTGCCATCGCCACCGAGGTCGCGACTATTGAGCCGACTACAGCCGATGTT
GTCGCTATCATTGTTGACCCGGGGCAGACGGCATGA

Fig. 3-11

>ORF12348c (SEQ ID NO:94)
CGGAAGGTGTATTCCTTGCTCAGTTCCTTTGTGAGGTACTTCGTCAGGCTTTCTGGCCAAGGCTTGTACTCGCCGGCATC
GGCCACCACCGCAGCCGCGTCTTCCGCGCCCGCCTCCACAGCTTGCTCCCCGGCGGCCTTGCGCTTGGCCGCAAGCGCGA
CAGCTGCGTTGTTGGCCTTCTCATGTAGTTCATGGAACTTCCAGACGAGCTCGTGGGCATCCCGAGGCTCGAACGCTACC
CGGATCCCGACGCACTTGCCGTCCCACTCAGTGCTCGGCAGCAGCGACATCGCAGTGGCGAGGTCGTCTTCACCGACGCG
GAACCAGAGATCCAAGAGTATGGACGGTAACCTTTTGGGCGCATCCTCATCGCCAGGGGGCGTCCTGCCGACCGCATCGA
TCTCGGCCCACAGCGCCGCACTGAAGTCAAAGAGCTCGAACTTCTTCACTTCGCCGCGAAGCATTGAGTACAGGCCTTGG
ACGGCGGATGTCTTCCCGCTGTTGTTGGCACCGACAAAGATAGAAATTTCGTCATCGAGCTCGATGACAACATCGCGCAG
CCGGCGGTAG

>ORF12314c (SEQ ID NO:96)
GGTACTTCGTCAGGCTTTCTGGCCAAGGCTTGTACTCGCCGGCATCGGCCACCACCGCAGCCGCGTCTTCCGCGCCCGCC
TCCACAGCTTGCTCCCCGGCGGCCTTGCGCTTGGCCGCAAGCGCGACAGCTGCGTTGTTGGCCTTCTCATGTAGTTCATG
GAACTTCCAGACGAGCTCGTGGGCATCCCGAGGCTCGAACGCTACCCGGATCCCGACGCACTTGCCGTCCCACTCAGTGC
TCGGCAGCAGCGACATCGCAGTGGCGAGGTCGTCTTCACCGACGCGGAACCAGAGATCCAAGAGTATGGACGGTAACCTT
TTGGGCGCATCCTCATCGCCAGGGGGCGTCCTGCCGACCGCATCGATCTCGGCCCACAGCGCCGCACTGA

>ORF13156c (SEQ ID NO:98)
CGACAGATTGCGCACATCCGTGTGATGGCCCAACTGGTCGTTGACGCGGCGGAAGTACCGAATGGGCGAGAATCCGCGTT
CATAGAGGATGTGCGGGGAGTGCGTGGTGATGACGAGCTGCGTGTGGAACAAAGTCGCGTGATCGTTAGCATCCTCAAGG
AGGCGCAAAACGTTCCTGATGAAGACCTGCTGGATCTGCGCGTGCAGATGCGCCTCAGGCTCCTCAATGAAGACCAAATG
AAGCGGAGCTCGCTTGTCATCCTCGGCTTTCCACTGCTCGTGCAAGTCGAGCAGCTCAACCACCATGTAGACCAGATTCT
TGAACCCCAGGCCATTGTAGCTGTCTGGCAGTTGGGCGGAAGCTACGCCCGGGATCACGTAGTGAACCTTGGCGTCTTGC
CCCAAGACAGTGGTCGGATCCAAGGCCGCCCGAATCACGATCTCCGGATTGTTGACGCCCGGATAGCCGAGCTTGGCCAG
GCGCGTGAGGGTGTCATTGAAGACTTCCTTCAGGTGGAAGTTCAGCTCCTTCTCCGAGGTATCTAG

>ORF12795 (SEQ ID NO:100)
CTTCCGCCCAACTGCCAGACAGCTACAATGGCCTGGGGGTTCAAGAATCTGGTCTACATGGTGGTTGAGCTGCTCGACTTG
CACGAGCAGTGGAAAGCCGAGGATGACAAGCGAGCTCCGCTTCATTTGGTCTTCATTGAGGAGCCTGAGGCGCATCTGCA
CGCGCAGATCCAGCAGGTCTTCATCAGGAACGTTTTGCGCCTCCTTGAGGATGCTAACGATCACGCGACTTTGTTCCACA
CGCAGCTCGTCATCACCACGCACTCCCCGCACATCCTCTATGAACGCGGATTCTCGCCCATTCGGTACTTCCGCCGCGTC
AACGACCAGTTGGGCCATCACACGGATGTGCGCAATCTGTCGCTATTCAAAACGGGCGCGTCCGACGCTCCAGCGCGCGA
ATTCCTGCAGCGGTATCTGA

>ORF12314c (SEQ ID NO:96)
GGTACTTCGTCAGGCTTTCTGGCCAAGGCTTGTACTCGCCGGCATCGGCCACCACCGCAGCCGCGTCTTCCGCGCCCGCC
TCCACAGCTTGCTCCCCGGCGGCCTTGCGCTTGGCCGCAAGCGCGACAGCTGCGTTGTTGGCCTTCTCATGTAGTTCATG
GAACTTCCAGACGAGCTCGTGGGCATCCCGAGGCTCGAACGCTACCCGGATCCCGACGCACTTGCCGTCCCACTCAGTGC
TCGGCAGCAGCGACATCGCAGTGGCGAGGTCGTCTTCACCGACGCGGAACCAGAGATCCAAGAGTATGGACGGTAACCTT
TTGGGCGCATCCTCATCGCCAGGGGGCGTCCTGCCGACCGCATCGATCTCGGCCCACAGCGCCGCACTGA

>ORF13156c (SEQ ID NO:98)
CGACAGATTGCGCACATCCGTGTGATGGCCCAACTGGTCGTTGACGCGGCGGAAGTACCGAATGGGCGAGAATCCGCGTT
CATAGAGGATGTGCGGGGAGTGCGTGGTGATGACGAGCTGCGTGTGGAACAAAGTCGCGTGATCGTTAGCATCCTCAAGG
AGGCGCAAAACGTTCCTGATGAAGACCTGCTGGATCTGCGCGTGCAGATGCGCCTCAGGCTCCTCAATGAAGACCAAATG
AAGCGGAGCTCGCTTGTCATCCTCGGCTTTCCACTGCTCGTGCAAGTCGAGCAGCTCAACCACCATGTAGACCAGATTCT
TGAACCCCAGGCCATTGTAGCTGTCTGGCAGTTGGGCGGAAGCTACGCCCGGGATCACGTAGTGAACCTTGGCGTCTTGC
CCCAAGACAGTGGTCGGATCCAAGGCCGCCCGAATCACGATCTCCGGATTGTTGACGCCCGGATAGCCGAGCTTGGCCAG
GCGCGTGAGGGTGTCATTGAAGACTTCCTTCAGGTGGAAGTTCAGCTCCTTCTCCGAGGTATCTAG

Fig. 3-12

>ORF12795 (SEQ ID NO:100)
CTTCCGCCCAACTGCCAGACAGCTACAATGGCCTGGGGGTTCAAGAATCTGGTCTACATGGTGGTTGAGCTGCTCGACTTG
CACGAGCAGTGGAAAGCCGAGGATGACAAGCGAGCTCCGCTTCATTTGGTCTTCATTGAGGAGCCTGAGGCGCATCTGCA
CGCGCAGATCCAGCAGGTCTTCATCAGGAACGTTTTGCGCCTCCTTGAGGATGCTAACGATCACGCGACTTTGTTCCACA
CGCAGCTCGTCATCACCACGCACTCCCCGCACATCCTCTATGAACGCGGATTCTCGCCCATTCGGTACTTCCGCCGCGTC
AACGACCAGTTGGGCCATCACACGGATGTGCGCAATCTGTCGCTATTCAAAACGGGCGCGTCCGACGCTCCAGCGCGCGA
ATTCCTGCAGCGGTATCTGA

>ORF13755c (SEQ ID NO:210)
GCTACCCGAACCCCAGCGCTGGAATCCTCAGCCAGCGACAGCGTCTTTTGCTCCGCCGTGACTTCCCAGAGCTCTGCCAT
CGACCGCTTCTTCGGGATCCAGCTGATGAGGGTTTGGTTGGACGTGACGGCACCTTCCACGTGTGCATGGCAGGTGCTGC
CACGCTTCTTGGACTTCTTCTTGCCACTCGGTTCTGCTTCGTCGTCGTCTTCAAGCTCGAAGGGCTTCAGGTCGTCGTCC
TCGTCCTCGTCATCTCCGTCAACGGCGCCCTCAGCGCCTGCGCCTTGCGCGGCGGCCTTCTCGGCGTCCGTCTTGACCGT
CACGCTGTCCAGATCCGTGATGACCAGTGTTGTGAGCCCAACGAAGGCGATCAGCTCCTGGAACCGATGCGCGAACGCAC
CACCGACTTCAAGGATGGTTAGGGCGGAAGAACGCAGGCGCTTGGCCACCAACTCGATCATTGCAGGCAGGAGCAGACGC
TCGACGTTGCCTTCCACCAATATCACCGCGTCGGAAAAAAAGAGATCGCAGTGCGTCAGCTTCAGATACCGCTGCAGGAA
TTCGCGCGCTGGAGCGTCGGACGCGCCCGTTTTGAATAGCGACAGATTGCGCACATCCGTGTGA

>ORF13795c (SEQ ID NO:212)
TGTCGTCGCACCCACCGTCACCGAAACCTTGGTCTGGTAAGCTACCCGAACCCCAGCGCTGGAATCCTCAGCCAGCGACA
GCGTCTTTTGCTCCGCCGTGACTTCCCAGAGCTCTGCCATCGACCGCTTCTTCGGGATCCAGCTGATGAGGGTTTGGTTG
GACGTGACGGCACCTTCCACGTGTGCATGGCAGGTGCTGCCACGCTTCTTGGACTTCTTCTTGCCACTCGGTTCTGCTTC
GTCGTCGTCTTCAAGCTCGAAGGGCTTCAGGTCGTCGTCCTCGTCCTCGTCATCTCCGTCAACGGCGCCCTCAGCGCCTG
CGCCTTGCGCGGCGGCCTTCTCGGCGTCCGTCTTGACCGTCACGCTGTCCAGATCCGTGATGACCAGTGTTGTGAGCCCA
ACGAAGGCGATCAGCTCCTGGAACCGATGCGCGAACGCACCACCGACTTCAAGGATGGTTAG

>ORF14727c (SEQ ID NO:214)
CAGGAAGTCGGCCGAGCTGAAGGATGTCCTCGTGGCCAAGTATGCCCTTGGCGTAGTCACTGCCCACGCCGTAGTTGAACG
TCCTGACGCCGGCCACAGCCTCCAGGCTTCGGACATATCGCTCTTGGTCGGCCTTGTTCCTGTCGCGCGTGGTCTGCCGG
ACACGCGAGCTGTAATTCTCGAACTCTTCTTCAAGTTCGGAGATCCGCCTGCGGATGTCGTTCTGCAGCCAAACCTTGAT
GTCGGCCTGGAACGTCTTTGCAATAGACCAGTAAAAGCTGTGGATGGTCGAGACATGAACCAGCGGGTCATCGTTGACGT
CCGCCAGGATTTCATTGGTGGCAAGGTCGGTATACGTGATGCACGCGACTATCTGCTTCCTCGCCCGCATGCTGGCGCCG
TGCTCCGAGATCACCCAGTCCAGCGCCTTGATGAGGGAGGTGGTCTTGCCGGAACCTGCGCCAGCACGAACCACGAAGGG
CTGCGGAGGCGTCGCTACAATGCATGCGTGGATCTCGCGGTCGGCGTCGGTATCTGGGCTATCAATTCGTCTGCTCATGC
CGTCTGCCCCGGGTCAACAATGATAGCGACAACATCGGCTGTAGTCGGCTCAATAGTCGCGACCTCGGTGGCGATGGCAG
CATCCGCCTCAAGCTCGTGGGCCACTTTGGCTTCGAGCCAGGCCAAGCCCTCGGCGATGTACGCGGGAACCTTCCAGCCA
TTGAGCGGCCCGCTTGCGAGTACCTCCAGCGCAAAGCGGGTCTTGTCGAAGTTCTTGCCGACCACCCTATCGTGTAACTT
CTCAGCCAGCTCTTCAGGGCTGCTCGGTGCGCGCTTGAGCTTGAGGCCGACCGACCGGTTTGCCTCAGCCTGGCACCAGT
CCGCGTTCTCAAGACCAAAGGCCTCCTCAAGTGTGCGGCCGCAGAGCTGTGATGTCGTCGCACCCACCGTCACCGAAACC
TTGGTCTGGTAA

>ORF13779 (SEQ ID NO:216)
CGGTGGGTGCGACGACATCACAGCTCTGCGGCCGCACACTTGAGGAGGCCTTTGGTCTTGAGAACGCGGACTGGTGCCAG
GCTGAGGCAAACCGGTCGGTCGGCCTCAAGCTCAAGCGCGCACCGAGCAGCCCTGAAGAGCTGGCTGAGAAGTTACACGA
TAGGGTGGTCGGCAAGAACTTCGACAAGACCCGCTTTGCGCTGGAGGTACTCGCAAGCGGGCCGCTCAATGGCTGGAAGG
TTCCCGCGTACATCGCCGAGGGCTTGGCCTGGCTCGAAGCCAAAGTGGCCCACGAGCTTGAGGCGGATGCTGCCATCGCC
ACCGAGGTCGCGACTATTGAGCCGACTACAGCCGATGTTGTCGCTATCATTGTTGACCCGGGGCAGACGGCATGAGCAGA
CGAATTGA

Fig. 3-13

>ORF14293c (SEQ ID NO:218)
GGGAGGTGGTCTTGCCGGAACCTGCGCCAGCACGAACCACGAAGGGCTGCGGAGGCGTCGCTACAATGCATGCGTGGATC
TCGCGGTCGGCGTCGGTATCTGGGCTATCAATTCGTCTGCTCATGCCGTCTGCCCCGGGTCAACAATGATAGCGACAACA
TCGGCTGTAGTCGGCTCAATAGTCGCGACCTCGGTGGCGATGGCAGCATCCGCCTCAAGCTCGTGGGCCACTTTGGCTTC
GAGCCAGGCCAAGCCCTCGGCGATGTACGCGGGAACCTTCCAGCCATTGAGCGGCCCGCTTGCGAGTACCTCCAGCGCAA
AGCGGGTCTTGTCGAAGTTCTTGCCGACCACCCTATCGTGTAA

>ORF14155 (SEQ ID NO:220)
CCCGGGGCAGACGGCATGAGCAGACGAATTGATAGCCCAGATACCGACGCCGACCGCGAGATCCACGCATGCATTGTAGC
GACGCCTCCGCAGCCCTTCGTGGTTCGTGCTGGCGCAGGTTCCGGCAAGACCACCTCCCTCATCAAGGCGCTGGACTGGG
TGATCTCGGAGCACGGCGCCAGCATGCGGGCGAGGAAGCAGATAGTCGCGTGCATCACGTATACCGACCTTGCCACCAAT
GAAATCCTGGCGGACGTCAACGATGACCCGCTGGTTCATGTCTCGACCATCCACAGCTTTTACTGGTCTATTGCAAAGAC
GTTCCAGGCCGACATCAAGGTTTGGCTGCAGAACGACATCCGCAGGCGGATCTCCGAACTTGAAGAAGAGTTCGAGAATT
ACAGCTCGCGTGTCCGGCAGACCACGCGCGACAGGAACAAGGCCGACCAAGAGCGATATGTCCGAAGCCTGGAGGCTGTG
GCCGGCGTCAGGACGTTCAACTACGGCGTGGGCAGTGACTACGCCAAGGGCATACTTGGCCACGAGGACATCCTTCAGCT
CGCCGACTTCCTGCTACAAAACCGCCCGCTGTTCCGACGGGTCGTGGCGCTGAGCTACCCGTTCGTGTTTATCGATGAGA
GTCAGGACACGTTCCCGGGTGTAGTGAAGTCTTTCAAGGAAGTGGAAGCCCAGATGCAGGGCAAGTTCTGCCTTGGTTTT
TTCGGCGACCCGATGCAGTCGATCTTCATGAGAGGCGCAGGGGACATCCAGCTTGAGGATCATTGGCGGGCCATCACGAA
GCCGGAGAACTTTCGCTGCGCCAAGCAGATCCTTGACGTCGCCAATGCCGTGCGCGCGCAGGGCGATGGCATGGAGCAAG
TCCGCGGGCTGCACGAGAGGGTCGATGGGAACCTCAAGCTGGTGGAGGGGTCGGCCCGGATGTTCGTCTTGCCGAACACG
CTGAACCGAACCGAGGCTTTGGCAAGAGTCCGAGCGTGGAGCTCGGCGACGAACAACGACGAGGGTTGGACAACCCCAGA
CATCGCAGTCAAGATTCTTGTCATCGTGCACCGCATGGCCGCAAACCGGCTTGGCTTCGGCGGCATCTACTCGGCGCTGA
ACGACAAGACGTCGGATGCCATGAAGCAAGGGATGCAGGACGGCACCGGTTGGCCCGTTCGACCCTTCCTAAGTTTTGCG
CTACCGATCGTTGCAGCTGTGAAGGCCGGCAATGAGTTCGCGGCGATGAGCCTGCTCCGGGAATTCAGCCCGCGCCTGGC
GCCTGCGGCTCTGACCGGCCGACGTGCCGCGGATGTATTGCGAGAGCTGCACGCTGCTGCGTCGAGGCTTGTCGCCATGC
TGGACGAGGCAGGGACCACCATTGGTGACATAGCTCTCCATCTCTGTGACACGGGTCTTTTTGAGTTCGACGAGCGCTAT
GCGCGTGTTCTTGGGTTTGTCAGGGATATTGCTGACACCGCTCAGGAGCCCGAGGCTGCTGATGCAGTTCCGGCCGAAGG
ATTATCCTTGGACGCGACAATGGCCAAGTTCTTCAATTGCTCTGCGCAAGAGCTTTGGCCCTATGAACGCTATGTCTCAG
AAGGCTCCCCCTATGCCACGCAGCACGGCGTGAAGGGAGCGCAGTTCGAACGCGTCATGGTGGTGATGGACGAGGAAGAA
AGCGACTACCGAACGTACAACTACGAGCGTGTCTTCGCGAGTGCTGAGGCCCGCGCTGCAGATCGTGCACGAGCACTAGA
CGGTGATGAAAACACTTGGAGCCGAACGCTGCGACTGCTTTACGTCTGCTGCACTCGTGCCCAGCGGGGGCTGGTACTAG
CGTTCTTTGTCGCCGACCCTGCGACCACCCTGGAAAACGTCGTGGCGAGCGGGATCTTGCCGCGAAGCGCAGTCTTTACG
CAGGAAGTGTTAGTTGGATGGCCATAG

>ORF14360 (SEQ ID NO:222)
TCGCGTGCATCACGTATACCGACCTTGCCACCAATGAAATCCTGGCGGACGTCAACGATGACCCGCTGGTTCATGTCTCG
ACCATCCACAGCTTTTACTGGTCTATTGCAAAGACGTTCCAGGCCGACATCAAGGTTTGGCTGCAGAACGACATCCGCAG
GCGGATCTCCGAACTTGAAGAAGAGTTCGAGAATTACAGCTCGCGTGTCCGGCAGACCACGCGCGACAGGAACAAGGCCG
ACCAAGAGCGATATGTCCGAAGCCTGGAGGCTGTGGCCGGCGTCAGGACGTTCAACTACGGCGTGGGCAGTGACTACGCC
AAGGGCATACTTGGCCACGAGGACATCCTTCAGCTCGCCGACTTCCTGCTACAAAACCGCCCGCTGTTCCGACGGGTCGT
GGCGCTGA

>ORF15342c (SEQ ID NO:224)
GAAGGGTCGAACGGGCCAACCGGTGCCGTCCTGCATCCCTTGCTTCATGGCATCCGACGTCTTGTCGTTCAGCGCCGAGT
AGATGCCGCCGAAGCCAAGCCGGTTTGCGGCCATGCGGTGCACGATGACAAGAATCTTGACTGCGATGTCTGGGGTTGTC
CAACCCTCGTCGTTGTTCGTCGCCGAGCTCCACGCTCGGACTCTTGCCAAAGCCTCGGTTCGGTTCAGCGTGTTCGGCAA
GACGAACATCCGGGCCGACCCCTCCACCAGCTTGAGGTTCCCATCGACCCTCTCGTGCAGCCCGCGGACTTGCTCCATGC
CATCGCCCTGCGCGCGCACGGCATTGGCGACGTCAAGGATCTGCTTGGCGCAGCGAAAGTTCTCCGGCTTCGTGATGGCC
CGCCAATGATCCTCAAGCTGGATGTCCCCTGCGCCTCTCATGAAGATCGACTGCATCGGGTCGCCGAAAAAACCAAGGCA
GAACTTGCCCTGCATCTGGGCTTCCACTTCCTTGAAAGACTTCACTACACCCGGGAACGTGTCCTGACTCTCATCGATAA
ACACGAACGGGTAGCTCAGCGCCACGACCCGTCGGAACAGCGGGCGGTTTTGTAG

Fig. 3-14

>ORF15260c (SEQ ID NO:226)
ATGCCGCCGAAGCCAAGCCGGTTTGCGGCCATGCGGTGCACGATGACAAGAATCTTGACTGCGATGTCTGGGGTTGTCCA
ACCCTCGTCGTTGTTCGTCGCCGAGCTCCACGCTCGGACTCTTGCCAAAGCCTCGGTTCGGTTCAGCGTGTTCGGCAAGA
CGAACATCCGGGCCGACCCCTCCACCAGCTTGAGGTTCCCATCGACCCTCTCGTGCAGCCCGCGGACTTGCTCCATGCCA
TCGCCCTGCGCGCGCACGGCATTGGCGACGTCAAGGATCTGCTTGGCGCAGCGAAAGTTCTCCGGCTTCGTGATGGCCCG
CCAATGA

>ORF14991 (SEQ ID NO:228)
CGTCGCCAATGCCGTGCGCGCGCAGGGCGATGGCATGGAGCAAGTCCGCGGGCTGCACGAGAGGGTCGATGGGAACCTCA
AGCTGGTGGAGGGGTCGGCCCGGATGTTCGTCTTGCCGAACACGCTGAACCGAACCGAGGCTTTGGCAAGAGTCCGAGCG
TGGAGCTCGGCGACGAACAACGACGAGGGTTGGACAACCCCAGACATCGCAGTCAAGATTCTTGTCATCGTGCACCGCAT
GGCCGCAAACCGGCTTGGCTTCGGCGGCATCTACTCGGCGCTGAACGACAAGACGTCGGATGCCATGAAGCAAGGGATGC
AGGACGGCACCGGTTGGCCCGTTCGACCCTTCCTAAGTTTTGCGCTACCGATCGTTGCAGCTGTGAAGGCCGGCAATGA

>ORF15590c (SEQ ID NO:230)
CGCTCGTCGAACTCAAAAAGACCCGTGTCACAGAGATGGAGAGCTATGTCACCAATGGTGGTCCCTGCCTCGTCCAGCAT
GGCGACAAGCCTCGACGCAGCAGCGTGCAGCTCTCGCAATACATCCGCGGCACGTCGGCCGGTCAGAGCCGCAGGCGCCA
GGCGCGGGCTGAATTCCCGGAGCAGGCTCATCGCCGCGAACTCATTGCCGGCCTTCACAGCTGCAACGATCGGTAGCGCA
AAACTTAGGAAGGGTCGAACGGGCCAACCGGTGCCGTCCTGCATCCCTTGCTTCATGGCATCCGACGTCTTGTCGTTCAG
CGCCGAGTAG

>ORF15675c (SEQ ID NO:232)
TCCTTCGGCCGGAACTGCATCAGCAGCCTCGGGCTCCTGAGCGGTGTCAGCAATATCCCTGACAAACCCAAGAACACGCG
CATAGCGCTCGTCGAACTCAAAAAGACCCGTGTCACAGAGATGGAGAGCTATGTCACCAATGGTGGTCCCTGCCTCGTCC
AGCATGGCGACAAGCCTCGACGCAGCAGCGTGCAGCTCTCGCAATACATCCGCGGCACGTCGGCCGGTCAGAGCCGCAGG
CGCCAGGCGCGGGCTGAATTCCCGGAGCAGGCTCATCGCCGCGAACTCATTGCCGGCCTTCACAGCTGCAACGATCGGTA
G

>ORF16405 (SEQ ID NO:234)
ATCGACTCTTTGAGGAAATGCGTGGGAAGCCTGGAAAAGTGCTGTTTCGCCTGCAAAGAAATAATTCATGTTCATGCGAT
TCGTTGTCGGCAGTGCGGCGAGTCCCAAGGCTGGCGAAGGTTCATGAGCTCTCCAACCTCAGTAGTTGCGTTGGTCCTTA
GCCTTTTATCAATCGCTGCCACAAAACCTGTGGAGCGATTGTTCGATGCCCAGCGAGCAGAGCTACAAATCTCCATCACG
GGTGGTGATTACAAAGCTGCCCAGCTTATGTTGACCAATAACGGGTCAAAGCCTGCAACTTTAGTTTCCTTCGAAATCAC
ATCGAAAGCCACGACCAATACGAAAACATGGTTTTTGGTAAGCAATACGGATGGCGAAATTCTGGAGCCAGGCAAAACTT
ACAAAATCAGGGCCTCAACCGATGAGTCTATCCCAAAAATTGTCGAAGCTGAGCGTCGGACGATTTTGAAGTCTCAGTAC
GCACTTGCAGATAATTGCGAATTAACCGCTAAATACATAGAGGCCACGGGGCAGAAGGTTGTGCGTGTGCAACCGTTCAT
GTGCGACACACCTCCTGAAAAGGGTGGCCTGCCCCCTGGTAAACCTGGCATACCCATTTGGTACCTTGGTCAAGAATGA

>ORF16925 (SEQ ID NO:236)
AGGCCACGGGGCAGAAGGTTGTGCGTGTGCAACCGTTCATGTGCGACACACCTCCTGAAAAGGGTGGCCTGCCCCCTGGT
AAACCTGGCATACCCATTTGGTACCTTGGTCAAGAATGATGTTTTTATGCCGCCCTGGGCTTTGACGCCGATTAAGCAAA
GCTGTGTTCGCTCATCCAATACGTCCCTCGCCCAGTTAAACGACTGTTATGTATATGGGTGCTGCCGCTACGTAATACCT
TGGCCCTACGCATACGAAGTTAATTCTGAAAGCGTTCAATGGACAATCTTCCTCCTCGGCGTCGACTGCAGCGGTAAGGT
GATCTACTTTCGAAACACTGCAAGGGTAGGTCCTTTTTTGGCAGCGTCCATATACCGACCGTGGTATGGCTCAGATGCGC
TGGTACTGCATTTCACCAAATAA

Fig. 3-15

>ORFl7793c (SEQ ID NO:238)
GCCAAAATGATTGTCATTGACAAAAATCTAGAACATCTTGTTGCGCAATGCGCTATATGTGAAAAAACTTTATTTGACGA
GTTTTCTCTCAAGATTCAATTGGGGCATACATATTACGAGCCAAAATCTTTGCCCGCCTCTGCAAGCATTGTATATGGGT
CGCATCCAGCCCCGTCGACGTTTTTTTTTGGAACCAAAAGAAATTCAGCAAAATTTGGTGCTGAAATCCGGTGAGCAAGTC
ATCACCTGCAGTAAACATCGATACAAAATACCGTTAGATTATTTTGGTCTGGTGCAAACCAAAGGAACCCTTGCGCGATT
GTTCGTGCAGGTAACCTGTAATGACGGTCAGGTAGAGCCGGGGTTCGACGGGTACGTAACCCTTGAAATCGTCAATATGT
CGCCTTGGACGATAGAAATACCGGCCGTGAGCGATATAGCACAACTTTATTTGGTGAAATGCAGTACCAGCGCATCTGAG
CCATACCACGGTCGGTATATGGACGCTGCCAAAAAAGGACCTACCCTTGCAGTGTTTCGAAAGTAG

>ORFl8548c (SEQ ID NO:240)
AGGACAATGGCAGGGTGGCCGCGTCTCGCAGCCCAAGGACGAAGGACAAATCTGATGAGTGTGTTACAGATCAAAGGGCG
TACAACGAAATCCCACACGGATTTTGACGCGGCATCGTACTCCAGCAACAGCCTTATACTCACTGATGCAGGGGACGAGA
GAATTGAAGAGTTTTCCCTCGAATTGTCCGTGGGTGAAGGGTGGAGTGATAACTATTCTGGCAACGACAAAAACCTGTGG
CGCATTGTCGATGGTATGACGATCAGGGGTCACGATTCTGTTGTGGTGGAGGCCGCTGAAGAAATCAAGGTGCCGCACAA
TCGGTACGGCATAGTCCTACCTACGGGAAGTCTTTTTCTCTCACGCGGCGTGCTGGTTGCTTCGGCGAAGGTCGAACCTG
CATTTGATGGCAAGCTCAAGCTCAGGATATTCAACACCACCAACAAAAATGTCTGCCTTACCAAAGGCGAGAAGCTTGGC
TCTGTGATTTTTTTCTCCACAGAATCGACGCACACCCAAAGCCCCATCAAGCGTGGCAGTGAAATATCGACGCTTCCCAT
CACGCGGCGCGCGCGATTGAAGAAGTGGTTTTCGCTCAATCCCACCATATGGGTCGGGTGGACGCTGAATTTAATCGGAA
GTTCCCTGGTGTCTTCTCTTATAATGTACGCCGTCTATTACAAGGTTGTGCTGGAACACCAGTCGCAGCCTCCTCAGTCA
CAACAAAACGCTCAGCCATCGCCGAACGAAGTTAAGCCAAAATGA

>ORFl7875 (SEQ ID NO:242)
ACGGCGTACATTATAAGAGAAGACACCAGGGAACTTCCGATTAAATTCAGCGTCCACCCGACCCATATGGTGGGATTGAG
CGAAAACCACTTCTTCAATCGCGCGCGCCGCGTGATGGGAAGCGTCGATATTTCACTGCCACGCTTGATGGGGCTTTGGG
TGTGCGTCGATTCTGTGGAGAAAAAAATCACAGAGCCAAGCTTCTCGCCTTTGGTAAGGCAGACATTTTTGTTGGTGGTG
TTGAATATCCTGAGCTTGAGCTTGCCATCAAATGCAGGTTCGACCTTCGCCGAAGCAACCAGCACGCCGCGTGAGAGAAA
AAGACTTCCCGTAGGTAGGACTATGCCGTACCGATTGTGCGGCACCTTGATTTCTTCAGCGGCCTCCACCACAACAGAAT
CGTGA

>ORFl8479 (SEQ ID NO:244)
TCTGTAACACACTCATCAGATTTGTCCTTCGTCCTTGGGCTGCGAGACGCGGCCACCCTGCCATTGTCCTTTATACCGGC
CGATATCCCCGGATACCGCCTGAAAGATGACGTGCGCAAAGCGTGCACCAATCTGAATTTCAAACGCCTCGCTGTGATTG
TTGGTGAGCGCGAACGTCATCGGCCCTACATAACCTGGAGGCAGCACACTGGAACTGAACGTTATCCCGCTTCTGAACAG
CGTGCTTCTCGGAAAAAAAAGCGCCGCCAGATCTTCCGGCAGATCGAATTCTTCCATGGTGCTCGCCAGATAAGTCTTGC
CCGGTTCCATGACGAAGCAGTCATCCGGGTCTGCGAGCACGACCTCGCTGGCAGGGGTGCGTCGCGTAGATTCTCGCAAG
CTTCCACCCCCTACTGTCAGGCGAGAGAGGCCTGCGAGTCTGAGGTCAAATCCAACGCCTTCCGGGGTGGTCAACTCACG
GTGGGCAAGGTGCTTGATTAG

>ORFl9027c (SEQ ID NO:246)
ATGATTTACTCACCGCACTCGCTCCTGAAACTGGTCCGGGATGGAAAACTAATCAAGCACCTTGCCCACCGTGAGTTGAC
CACCCCGGAAGGCGTTGGATTTGACCTCAGACTCGCAGGCCTCTCTCGCCTGACAGTAGGGGGTGGAAGCTTGCGAGAAT
CTACGCGACGCACCCCTGCCAGCGAGGTCGTGCTCGCAGACCCGGATGACTGCTTCGTCATGGAACGGGGCAAGACTTAT
CTGGCGAGCACCATGGAAGAATTCGATCTGCCGGAAGATCTGGCGGCGCTTTTTTTTTCCGAGAAGCACGCTGTTCAGAAG
CGGGATAACGTTCAGTTCCAGTGTGCTGCCTCCAGGTTATGTAGGGCCGATGACGTTCGCGCTCACCAACAATCACAGCG
AGGCGTTTGAAATTCAGATTGGTGCACGCTTTGCGCACGTCATCTTTCAGGCGGTATCCGGGGATATCGGCCGGTATAAA
GGACAATGGCAGGGTGGCCGCGTCTCGCAGCCCAAGGACGAAGGACAAATCTGA

>ORFl9305 (SEQ ID NO:248)
TGGCCGTTCTCTGCCTGTCGCCTCTTTGGCATGACTGGTCAAGTCGGATGCAAACGGTGGTCAGCACCAATGCAATTGGG
TGGTCATGTGCGATGCAATTACGCAGTTGAGCCTGGCCCAGTTCCTCCCAAGCAAAGCATAAGACCAAGATGGCACATTG
CCAACAAAATACCCTTCCCCGCTACCGTTGTTTTATCGTTGTTGCCAGCCCTGATCTGGCGGAAAAGCCCGCTCCATGAA
TCGTCATGGAGCCTCCCATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCCGTCCCCCACCCCAACAACCAAAGCTGCC
CCAGGGGGGATTCATCCTTCCTCTGA

<p style="text-align:center">Fig. 3-16</p>

>ORF19519 (SEQ ID NO:250)
TCTGGCGGAAAAGCCCGCTCCATGAATCGTCATGGAGCCTCCCATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCCGT
CCCCCACCCCAACAACCAAAGCTGCCCCAGGGGGATTCATCCTTCCTCTGAGCAGCATGGAACTGCTCGGCACGCCTCGC
CGCCGGCAGCTACTGGAGAACATCTGGCAGCGCGCCTCGCTATCCAAGCAGCAATTCGAGGAGATCTACCGGCGGCCACT
GGCCAACTATGCCGAGCTGGTCCAGCAGCTCCCTGCTTCGGAAAATCATCACCATGCCCATCCAGGCGGGATGATCGATC
ACGGCCTGGAGATCGTGGCCTACGCACTCAAGGTACGGCAGACCTACCTGCTCCCGATCGGCGCAGCGCCGGAGTCACAG
TCAGCCCAGGCTGA

>ORF19544 (SEQ ID NO:252)
ATCGTCATGGAGCCTCCCATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCCGTCCCCCACCCCAACAACCAAAGCTGC
CCCAGGGGGATTCATCCTTCCTCTGAGCAGCATGGAACTGCTCGGCACGCCTCGCCGCCGGCAGCTACTGGAGAACATCT
GGCAGCGCGCCTCGCTATCCAAGCAGCAATTCGAGGAGATCTACCGGCGGCCACTGGCCAACTATGCCGAGCTGGTCCAG
CAGCTCCCTGCTTCGGAAAATCATCACCATGCCCATCCAGGCGGGATGATCGATCACGGCCTGGAGATCGTGGCCTACGC
ACTCAAGGTACGGCAGACCTACCTGCTCCCGATCGGCGCAGCGCCGGAGTCACAGTCAGCCCAGGCTGAAGCCTGGTCGG
CCGCCGCGGCGTATGGCGCCCTGGCTCATGACATAGGCAAGATCGTCGTCGACCTGCAGGTTGAGCTACAGGACGGCAGC
ACCTGGCACCCTTGGAACGGACCGATCAACCAGCCATACCGCTTCAAGTACGTGAAGTCCCGCGAATACCAGCTCCACGG
CGCTGCCTCAGCACTTCTCATCCACCAACTGCTACCGCGCACTGCACTCGATTGGCTCAGTCGCTTTCCAGAGCTGTGGG
CTCAATTGATCTACCTGTTCGCTGGGCAGTACGAGCACGCCGGGATCCTCGGCGAGATCATCGTGAAGGCAGACCAGGCC
TCAGTTGCACAGGAGCTAGGAGGCAATCCGGATCGAGCTCTGGCTGCACCGAAGCAGTCGCTGCAGCGGCAGTTGGCAGA
CGGCCTTCGCTTCTTGGTGAAGGACAAGTTCAAGTTGAATCAACCTAGCGGCCCGTCTGATGGATGGCTGACCCAGGACG
CACTCTGGCTGGTGAGCAAGCCTGCTGCCGATCAACTGAGAGCCTACCTGCTGGCCCAGGGTATCGATGGGGTGCCCTCC
TCTAACGCGCCGTTCTTCAGCATGCTCCAGGACCAAGCCGTCATCCAGACAAATGCCGAGGACAAGGCCATTTGGACGGC
CACGGTAGACAACGGTGCTGGATGGAGAAACAAGTTCACGCTACTCAAGATTGCTCCAGCCTTGATCTGGACAGATGCTG
CCGAGCGCCCCTCACCCTACAGCGGATCACTGGTCGTTGAAGATGGAACCGCCTCAACGGAAAAGCCGGAAACGACCTGT
GAAATTCCCAACGGGCCGGCTGAACAGCAGCAAGCACCAGAAACGAAGATGATGCTCCATCAACCTGCGCCGAGCGTTGC
GAAACCGGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGATGATCAAGAAGAAACAGACGATTTGTATGCACTTC
TTGGTAATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACGACTCGCCGGCTGCCTCTCCTACGAACACACGCGGG
GAGGAGAACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGATTGCGCTCCTGAAGCAATTGAAGATGTATTTATGCC
TAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATCTGGCATCGCGGCCCGTCGCCTGTTCATCAACG
ACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAG
CATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGG
GCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGG
CCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCC
GAAGGAGGTGTGGAATGA

>ORF20008 (SEQ ID NO:254)
GCTACAGGACGGCAGCACCTGGCACCCTTGGAACGGACCGATCAACCAGCCATACCGCTTCAAGTACGTGAAGTCCCGCG
AATACCAGCTCCACGGCGCTGCCTCAGCACTTCTCATCCACCAACTGCTACCGCGCACTGCACTCGATTGGCTCAGTCGC
TTTCCAGAGCTGTGGGCTCAATTGATCTACCTGTTCGCTGGGCAGTACGAGCACGCCGGGATCCTCGGCGAGATCATCGT
GAAGGCAGACCAGGCCTCAGTTGCACAGGAGCTAGGAGGCAATCCGGATCGAGCTCTGGCTGCACCGAAGCAGTCGCTGC
AGCGGCAGTTGGCAGACGGCCTTCGCTTCTTGGTGAAGGACAAGTTCAAGTTGAATCAACCTAG

>ORF20623c (SEQ ID NO:256)
CGTGAACTTGTTTCTCCATCCAGCACCGTTGTCTACCGTGGCCGTCCAAATGGCCTTGTCCTCGGCATTTGTCTGGATGA
CGGCTTGGTCCTGGAGCATGCTGAAGAACGGCGCGTTAGAGGAGGGCACCCCATCGATACCCTGGGCCAGCAGGTAGGCT
CTCAGTTGATCGGCAGCAGGCTTGCTCACCAGCCAGAGTGCGTCCTGGGTCAGCCATCCATCAGACGGGCCGCTAGGTTG
ATTCAACTTGAACTTGTCCTTCACCAAGAAGCGAAGGCCGTCTGCCAACTGCCGCTGCAGCGACTGCTTCGGTGCAGCCA
GAGCTCGATCCGGATTGCCTCCTAG

Fig. 3-17

>ORF21210c (SEQ ID NO:258)
CGCTTGAAAATTCCTGGCGTGACCAGCATGGCGGTCCCGTCTACGGTATGCACCAAAGCCTTGGTGTCGTTGATGAACAG
GCGACGGGCCGCGATGCCAGATTTCATCCAACCAACGAATCCCTGTCCCAGATCAGTACTTCTGCTAGGCATAAATACAT
CTTCAATTGCTTCAGGAGCGCAATCTGTTGGCTCCTTGGTCCCTAGTGGCTGCTGTAGGTTCTCCTCCCCGCGTGTGTTC
GTAGGAGAGGCAGCCGGCGAGTCGTGGCTAGTGTCTAGCTCTTCTAGTGGCGAATTGATATTACCAAGAAGTGCATACAA
ATCGTCTGTTTCTTCTTGA

>ORF21493c (SEQ ID NO:260)
GCTGCTGCGGCGTCATTCCACACCTCCTTCGGCATCGGTGATGACCGTGAGGCTTGGGTTGTCCAGAGGCTGCTCAGGGA
ACAGCAATTTGGGATCCTGGAGCAGGTAGGCCTTGAGCTCTTTCGTCTTGCGAGGACCAGAAACCTTGATGGTCCAGATG
TTCAGGTTTTTACTGGTCTTCCGATGAAGCCCCTGTTTTTCGAACGCGCGCTGCACCAGCTTCCAGCCGGTCGTCTCCTT
GGCTTGGGCCAGTTTTTCAAGCACCGGATGCTCTTGGACATAGCGCTTGAAAATTCCTGGCGTGACCAGCATGGCGGTCC
CGTCTACGGTATGCACCAAAGCCTTGGTGTCGTTGATGAACAGGCGACGGGCCGCGATGCCAGATTTCATCCAACCAACG
AATCCCTGTCCCAGATCAGTACTTCTGCTAGGCATAAATACATCTTCAATTGCTTCAGGAGCGCAATCTGTTGGCTCCTT
GGTCCCTAG

>ORF21333 (SEQ ID NO:262)
ACATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTG
TTCCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGACGCCGCAGCAG
CTCACCGAGGAGTACATCTTCGCGCACGATCTCCGGGAAGCCAGCGCGAAGATCTACCGCGCCGCGACCAAGGCGCTGCT
CAAGCACTTCGGTCCTACGGCAACCGTACAGGACGTGGACCACCGGGCTGTCCTGGGATGGCGACGCAAGGTACTGGAAC
AAGGCCTGTCCAAGCGGAGCTGGAACACGTACTCGAATCATCTGCGGACGATCTGGGGCTATGCCATCGAGCATGA

>ORF22074c (SEQ ID NO:264)
GTGAGGCACAAGCCCCTCCGTTATTGGCACTACGAACTCTTTGTGAGTCTTCTCTGTCTCGCCGCGGATGAGGATCAGTT
GATTTTCCCAGTCGATGTCGCGCTTGCGGATGCACAACAGCGCATTCAACCGGATGCCGGTGAAGTAGAAGACCTCAAAC
GTGCAAAGCCAGAACCAGGCGGGCGTGATCCGTGCGCGTTCGCCGGTGCAGCGCTCTGCGCCGTCCTGCATGTTGAGCCA
ATTGCGGGCGAGCAGGATGGCTTCGGCGGCGACGGTTTTGCTTGCTCGCCTGGGGGGAATGACGGTGGTCTTTCTGAACG
GGTTGACTTGGGAGTGCGTCACCAACTCATGCTCGATGGCATAGCCCCAGATCGTCCGCAGATGATTCGAGTACGTGTTC
CAGCTCCGCTTGGACAGGCCTTGTTCCAGTACCTTGCGTCGCCATCCCAGGACAGCCCGGTGGTCCACGTCCTGTACGGT
TGCCGTAGGACCGAAGTGCTTGAGCAGCGCCTTGGTCGCGGCGCGGTAGATCTTCGCGCTGGCTTCCCGGAGATCGTGCG
CGAAGATGTACTCCTCGGTGAGCTGCTGCGGCGTCATTCCACACCTCCTTCGGCATCGGTGATGACCGTGAGGCTTGGGT
TGTCCAGAGGCTGCTCAGGGAACAGCAATTTGGGATCCTGGAGCAGGTAG

>ORF21421 (SEQ ID NO:266)
GCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGACGCCGCAGCAGCTCACCGA
GGAGTACATCTTCGCGCACGATCTCCGGGAAGCCAGCGCGAAGATCTACCGCGCCGCGACCAAGGCGCTGCTCAAGCACT
TCGGTCCTACGGCAACCGTACAGGACGTGGACCACCGGGCTGTCCTGGGATGGCGACGCAAGGTACTGGAACAAGGCCTG
TCCAAGCGGAGCTGGAACACGTACTCGAATCATCTGCGGACGATCTGGGGCTATGCCATCGAGCATGAGTTGGTGACGCA
CTCCCAAGTCAACCCGTTCAGAAAGACCACCGTCATTCCCCCCAGGCGAGCAAGCAAAACCGTCGCCGCCGAAGCCATCC
TGCTCGCCCGCAATTGGCTCAACATGCAGGACGGCGCAGAGCGCTGCACCGGCGAACGCGCACGGATCACGCCCGCCTGG
TTCTGGCTTTGCACGTTTGAGGTCTTCTACTTCACCGGCATCCGGTTGAATGCGCTGTTGTGCATCCGCAAGCGCGACAT
CGACTGGGAAAATCAACTGATCCTCATCCGCGGCGAGACAGAGAAGACTCACAAAGAGTTCGTAGTGCCAATAACGGAGG
GGCTTGTGCCTCACCTATCGAGGCTCCTGCAAGAGGCCGATAGAGCCGGATTCGCCGATGACGACCAGTTGTTCAACGTC
AACCGGTTCTCACCGCACTACAAGAGCAAGGTGATGAACTCCGACCAGGTCGAAGCCATGTACCGGAAGTTGACCGAGAA
GGTTGGGGTGCGGATGACCCCGCACCGTTTCCGGCACACCCTGGCCACCGACTTGATGAAGGCACCCGAGCGGAACATTC
ACCTCACGAAGTGCCTGCTCAACCACTCGAATATCCAGACCACGATGAGCTACATCGAGGCCGACTACGATCACATGCGT
GCCGTGCTGCATGCTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGTCAGGAAGGTGGATTACAGCGGCTCCCCGCAAGC
CTCTGCCAAACCGAAGCCATGCGGGCAACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAGGCCAGGACAGAGCCTG
CAGAACCAAGGGAGCACACGCCAGGGACAGGCATTCAGGGAGGTCCAACCGCGTGGGAAGCAGATGCGCTACCACAGCCA
CCTGACACCTTCGAACCAAGCGTGCTGTTCACTCTGATGGCTCAAAACTTATCGAACCGTGCCGCCTCGGCATCCGCGGC
TCCCGCTGCAACAAGCGGATCAGGCGGATGGGGATCTGCCGCCCGAAGCAATCTCGCCTAG

Fig. 3-18

>ORF22074c (SEQ ID NO:264)
GTGAGGCACAAGCCCCTCCGTTATTGGCACTACGAACTCTTTGTGAGTCTTCTCTGTCTCGCCGCGGATGAGGATCAGTT
GATTTTCCCAGTCGATGTCGCGCTTGCGGATGCACAACAGCGCATTCAACCGGATGCCGGTGAAGTAGAAGACCTCAAAC
GTGCAAAGCCAGAACCAGGCGGGCGTGATCCGTGCGCGTTCGCCGGTGCAGCGCTCTGCGCCGTCCTGCATGTTGAGCCA
ATTGCGGGCGAGCAGGATGGCTTCGGCGGCGACGGTTTTGCTTGCTCGCCTGGGGGGAATGACGGTGGTCTTTCTGAACG
GGTTGACTTGGGAGTGCGTCACCAACTCATGCTCGATGGCATAGCCCCAGATCGTCCGCAGATGATTCGAGTACGTGTTC
CAGCTCCGCTTGGACAGGCCTTGTTCCAGTACCTTGCGTCGCCATCCCAGGACAGCCCGGTGGTCCACGTCCTGTACGGT
TGCCGTAGGACCGAAGTGCTTGAGCAGCGCCTTGGTCGCGGCGCGGTAGATCTTCGCGCTGGCTTCCCGGAGATCGTGCG
CGAAGATGTACTCCTCGGTGAGCTGCTGCGGCGTCATTCCACACCTCCTTCGGCATCGGTGATGACCGTGAGGCTTGGGT
TGTCCAGAGGCTGCTCAGGGAACAGCAATTTGGGATCCTGGAGCAGGTAG

>ORF21421 (SEQ ID NO:266)
GCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGACGCCGCAGCAGCTCACCGA
GGAGTACATCTTCGCGCACGATCTCCGGGAAGCCAGCGCGAAGATCTACCGCGCCGCGACCAAGGCGCTGCTCAAGCACT
TCGGTCCTACGGCAACCGTACAGGACGTGGACCACCGGGCTGTCCTGGGATGGCGACGCAAGGTACTGGAACAAGGCCTG
TCCAAGCGGAGCTGGAACACGTACTCGAATCATCTGCGGACGATCTGGGGCTATGCCATCGAGCATGAGTTGGTGACGCA
CTCCCAAGTCAACCCGTTCAGAAAGACCACCGTCATTCCCCCCAGGCGAGCAAGCAAAACCGTCGCCGCCGAAGCCATCC
TGCTCGCCCGCAATTGGCTCAACATGCAGGACGGCGCAGAGCGCTGCACCGGCGAACGCGCACGGATCACGCCCGCCTGG
TTCTGGCTTTGCACGTTTGAGGTCTTCTACTTCACCGGCATCCGGTTGAATGCGCTGTTGTGCATCCGCAAGCGCGACAT
CGACTGGGAAAATCAACTGATCCTCATCCGCGGCGAGACAGAGAAGACTCACAAAGAGTTCGTAGTGCCAATAACGGAGG
GGCTTGTGCCTCACCTATCGAGGCTCCTGCAAGAGGCCGATAGAGCCGGATTCGCCGATGACGACCAGTTGTTCAACGTC
AACCGGTTCTCACCGCACTACAAGAGCAAGGTGATGAACTCCGACCAGGTCGAAGCCATGTACCGGAAGTTGACCGAGAA
GGTTGGGGTGCGGATGACCCCGCACCGTTTCCGGCACACCCTGGCCACCGACTTGATGAAGGCACCCGAGCGGAACATTC
ACCTCACGAAGTGCCTGCTCAACCACTCGAATATCCAGACCACGATGAGCTACATCGAGGCCGACTACGATCACATGCGT
GCCGTGCTGCATGCTAGAAGCCTGGCCCAAGGCGCGCTGGAGAATGTCAGGAAGGTGGATTACAGCGGCTCCCCGCAAGC
CTCTGCCAAACCGAAGCCATGCGGGCAACCTCTCGCTCGAGTGAGTGAAGCGCCGCCACCGGAGGCCAGGACAGAGCCTG
CAGAACCAAGGGAGCACACGCCAGGGACAGGCATTCAGGGAGGTCCAACCGCGTGGGAAGCAGATGCGCTACCACAGCCA
CCTGACACCTTCGAACCAAGCGTGCTGTTCACTCTGATGGCTCAAAAACTTATCGAACCGTGCCGCCTCGGCATCCGCGGC
TCCCGCTGCAACAAGCGGATCAGGCGGATGGGGATCTGCCGCCCGAAGCAATCTCGCCTAG

>ORF22608c (SEQ ID NO:268)
CGCATCTGCTTCCCACGCGGTTGGACCTCCCTGAATGCCTGTCCCTGGCGTGTGCTCCCTTGGTTCTGCAGGCTCTGTCC
TGGCCTCCGGTGGCGGCGCTTCACTCACTCGAGCGAGAGGTTGCCCGCATGGCTTCGGTTTGGCAGAGGCTTGCGGGGAG
CCGCTGTAATCCACCTTCCTGACATTCTCCAGCGCGCCTTGGGCCAGGCTTCTAGCATGCAGCACGGCACGCATGTGATC
GTAGTCGGCCTCGATGTAGCTCATCGTGGTCTGGATATTCGAGTGGTTGAGCAGGCACTTCGTGAGGTGAATGTTCCGCT
CGGGTGCCTTCATCAAGTCGGTGGCCAGGGTGTGCCGGAAACGGTGCGGGGTCATCCGCACCCCAACCTTCTCGGTCAAC
TTCCGGTACATGGCTTCGACCTGGTCGGAGTTCATCACCTTGCTCTTGTAGTGCGGTGA

>ORF22626 (SEQ ID NO:270)
CACCTTCGAACCAAGCGTGCTGTTCACTCTGATGGCTCAAAAACTTATCGAACCGTGCCGCCTCGGCATCCGCGGCTCCCG
CTGCAACAAGCGGATCAGGCGGATGGGGATCTGCCGCCCGAAGCAATCTCGCCTAGCGATACCGGTACTGAGGGCCGGCT
ACCGGACGAAAGGTAGCCGTGCCTTCCAGCAGATCGTTAGGCCTGTAGGAAAAATCTGGAATTACCGAGAGCGCCTGGAT
TCCAGCGCCGGCATGCTGGCAGAGCCAGCGCAATTTCAAGGCCAATACCACAGTACCCTCTGTAATCGCTGA

Fig. 3-19

>ORF23228 (SEQ ID NO:272)
AGAGATTCGAACTCCCGACATCCTGCTCCCAAAGCAGGCGCGCTACCGGACTGCGCTATACCCCGATTGGAATTTGGCTC
CGCGACCTGGACTCGAACCAGGGACCCAATGATTAACAGTCATTTGCTCTACCGACTGAGCTATCGCGGAACGTCTTTCT
TCCAACCCTGGACGCTTCCGGTGTTGCTGGATTCGCGTCTCAGAGGCGCGCCATTTTACGGATGCGCGCGGGCATGTCAA
CCCTCTGATCCAAAAAGTTTTTCTTCTTTTTCCACGAGCGACAAAACGGCCCTTCCACTGCATGCGGCAGCGCTCTCGCG
CCTACCGGACGCCCATGAAAAAGCCCCGCCGAAGCGGGGCTTTCCCTGTCCGCCCCCGAAGAGGTCAGGCGAAGACGATC
TCGTCGCCTTCCACCTTCGCCGAGATACTGGCACCCGGCGCGAATTTGCCGGCCAGGATCAGTTGCGCCAGCGGGTTCTC
GATCCAGCGCTGGATGGCCCGCTTCAGCGGGCGTGCGCCATAGACCGGGTCGAAGCCGACGGCAATCAGCTTGTCCAGCG
CCTCCTGGCTCAGTTCCAGGCTCAGCTCGCGCTCGGCCAGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGGCGATGCCG
GCGATCTGCTCGCGAGCCAGCGGCTCGAACACCACCACTTCGTCGATCCGGTTGATGAATTCCGGACGGAAGTGCGCATT
GACCGCGTCCATCACTGCGGCACGTTGCGCCTCGCGGTCGCCGGCCAGCTCCTGGATCTGCGCCGAACCGAGGTTGGAGG
TCATCACCACCACGGTGTTGCGGAAGTCCACCGTACGCCCGTGACTGTCGGTCAGGCGTCCGTCCTCGAGCACCTGGAGG
AGAATGTTGAATACATCCGGATGGGCCTTCTCCACCTCGTCCAGCAGCACCACCGAGTAGGGCTTGCGGCGGATCGCCTC
GGTCAGGTAGCCGCCTTCCTCGAAGCCGACGTAGCCCGGAGGCGCGCCGATCAGGCGGGCCACCGAGTGTTTCTCCATGA
ACTCGGACATATCTATCCGCACCAGCGCCTCCTCGGTATCGAAGAGGAACTCGGCCAGCGCCTTGCACAACTCGGTCTTG
CCCACCCCGGTCGGGCCGAGGAAGAGGAACGAGCCGCTCGGCCGGTTCGGATCGGCGAGGCCGGCGCGCGAACGGCGCAC
GGCGTTGGACACGGCGACTACCGCCTCGTCCTGGCCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGCTTCT
CGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTGGAAACCACTTCGGCGATTTCCTCGTCGGTCACC
TTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGTGCTGGTCGACCATCTGCAGGCTGCGTTCCAGGTCCGGGATGGTCTG
GTACTGGATGCGCGCCATGCTCTCGAGGTCGCCCTTGCGCCGCGCCGCCTCCATCTCCTGCTTGGCCTGCTCGATCTTCT
GCTGGATCTGCGCCGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAGGTCGGCGTATTCGCGCTCGAGC
TTGACGATATCCTCCTCCAGCTTGGCCAGGCGCTTCCTGGTGGCTTCGTCGTCTTCCTTCTTCAGCGCCTCGCGCTCGAT
CTTCAGCTGGATCAGGCGACGGTCGAGACGATCCAGTTCCTCCGGCTTGGAGTCGATCTCCATGCGGATGCGGCTGGCGG
CCTCGTCGATCAGGTCGATGGCCTTGTCCGGCAGTTGCCGATCGGTGATGTAGCGGTGCGACAGCTTGGCCGCGGCGATG
ATCGCGCCGTCGGTGATGCTCACCCCGTGGTGCACTTCATAGCGTTCCTTGAGGCCACGGAGGATGGCGATGGTGTCTTC
CTCGCTCGGTTCGTCCACCAGCACCTTCTGGAAGCGGCGCTCCAGCGCGGCATCCTTCTCGATGTACTGGCGATACTCGT
CGAGGGTAGTAGCACCGACGCAGTGCAGCTCGCCGCGCGCCAGAGCCGGCTTGAGCATGTTGCCGGCGTCCATGGCACCT
TCCGCCTTGCCGGCGCCGACCATGGTGTGCAGTTCGTCGATGAACAGGATGACCCGGCCTTCCTGCTTGCCCAGTTCGTT
GAGGACCGCCTTCAGGCGTTCCTCGAACTCGCCGCGGAACTTGGCACCGGCGATCAGCGCCCCCATGTCCAGGGCCAGCA
GGCGCTTGTCCTTGAGGCCGTCCGGCACTTCGCCGTTGATGATGCGCTGGGCCAGGCCCTCGACGATGGCGGTCTTGCCG
ACGCCGGGTTCGCCGATCAGCACCGGGTTGTTCTTGGTCCGCCGCTGCAGGACCTGGATGGTCCGGCGGATCTCGTCGTC
GCGACCGATCACCGGGTCGAGCTTGCCTTCCTCGGCGCGCTTGGTCATGTCGACGGTGTACTTGTCCAGCGCCTGGCGCG
ACTCCTCGACGTTCGGGTCGTTCACCGCTTCGCCGCCACGCAGGTTGGCCACGGCATTCTCCAGCGCCTTGCGCGACACG
CCCTGGCCGAGCAGCAGCTTGCCGAGCCTGGTGTTCTCGTCCATCGCGGCCAGCAATACCAGCTCGCTGGAGATGAACTG
GTCGCCCTTCTGCTGGGCCAGGCGGTCAGCCTGGTTGAGCAGGCGTGCGAGATCCTGGGACAGGTTCACGTCGCCGGTCG
GGCTCTGGATCTTCGGCAGCGCGTCGAGTTCTTTGTTGAGGCCGCTGCGCAGGGCGGCGATATCGAAGCCGACCTGCATC
AGCAGGGGCTTGATCGAACCGCCTTGCTGCTCGAGCAGGGCGGAAAGCAGGTGCACCGGCTCGATGGCCGGATGGTCATG
GCCAACGGCCAGGGACTGGGCGTCGGAGAGCGCCAGTTGCAGCTTGCTGGTCAAACGGTCTATTCGCATGGGTCGTCCTT
CCTTCTATAG

>ORF23367 (SEQ ID NO:274)
GCTATCGCGGAACGTCTTTCTTCCAACCCTGGACGCTTCCGGTGTTGCTGGATTCGCGTCTCAGAGGCGCGCCATTTTAC
GGATGCGCGCGGGCATGTCAACCCTCTGATCCAAAAAGTTTTTCTTCTTTTTCCACGAGCGACAAAACGGCCCTTCCACT
GCATGCGGCAGCGCTCTCGCGCCTACCGGACGCCCATGAAAAAGCCCCGCCGAAGCGGGGCTTTCCCTGTCCGCCCCCGA
AGAGGTCAGGCGAAGACGATCTCGTCGCCTTCCACCTTCGCCGAGATACTGGCACCCGGCGCGAATTTGCCGGCCAGGAT
CAGTTGCGCCAGCGGGTTCTCGATCCAGCGCTGGATGGCCCGCTTCAGCGGGCGTGCGCCATAG

Fig. 3-20

>ORF25103c (SEQ ID NO:276)
AGTGCACCACGGGGTGAGCATCACCGACGGCGCGATCATCGCCGCGGCCAAGCTGTCGCACCGCTACATCACCGATCGGC
AACTGCCGGACAAGGCCATCGACCTGATCGACGAGGCCGCCAGCCGCATCCGCATGGAGATCGACTCCAAGCCGGAGGAA
CTGGATCGTCTCGACCGTCGCCTGATCCAGCTGAAGATCGAGCGCGAGGCGCTGAAGAAGGAAGACGACGAAGCCACCAG
GAAGCGCCTGGCCAAGCTGGAGGAGGATATCGTCAAGCTCGAGCGCGAATACGCCGACCTCGAGGAGATCTGGAAGTCCG
AGAAGGCCGAGGTGCAGGGCTCGGCGCAGATCCAGCAGAAGATCGAGCAGGCCAAGCAGGAGATGGAGGCGGCGCGCGGCGC
AAGGGCGACCTCGAGAGCATGGCGCGCATCCAGTACCAGACCATCCCGGACCTGGAACGCAGCCTGCAGATGGTCGACCA
GCACGGCAAGACCGAGAACCAGTTGCTGCGCAACAAGGTGACCGACGAGGAAATCGCCGAAGTGGTTTCCAAGTGGACCG
GTATCCCGGTGTCGAAGATGCTCGAGGGCGAGCGCGAGAAGCTGCTGCGCATGGAGCAGGAGCTGCATCGGCGAGTGATC
GGCCAGGACGAGGCGGTAGTCGCCGTGTCCAACGCCGTGCGCCGTTCGCGCGCCGGCCTCGCCGATCCGAACCGGCCGAG
CGGCTCGTTCCTCTTCCTCGGCCCGACCGGGGGTGGGCAAGACCGAGTTGTGCAAGGCGCTGGCCGAGTTCCTCTTCGATA
CCGAGGAGGCGCTGGTGCGGATAGATATGTCCGAGTTCATGGAGAAACACTCGGTGGCCCGCCTGATCGGCGCGCCTCCG
GGCTACGTCGGCTTCGAGGAAGGCGGCTACCTGACCGAGGCGATCCGCCGCAAGCCCTACTCGGTGGTGCTGCTGGACGA
GGTGGAGAAGGCCCATCCGGATGTATTCAACATTCTCCTCCAGGTGCTCGAGGACGGACGCCTGACCGACAGTCACGGGC
GTACGGTGGACTTCCGCAACACCGTGGTGGTGATGACCTCCAACCTCGGTTCGGCGCAGATCCAGGAGCTGGCCGGCGAC
CGCGAGGCGCAACGTGCCGCAGTGATGGACGCGGTCAATGCGCACTTCCGTCCGGAATTCATCAACCGGATCGACGAAGT
GGTGGTGTTCGAGCCGCTGGCTCGCGAGCAGATCGCCGGCATCGCCGAGATCCAGCTCGGTCGCCTGCGCAAGCGCCTGG
CCGAGCGCGAGCTGAGCCTGGAACTGAGCCAGGAGGCGCTGGACAAGCTGATTGCCGTCGGCTTCGACCCGGTCTATGGC
GCACGCCCGCTGAAGCGGGCCATCCAGCGCTGGATCGAGAACCCGCTGGCGCAACTGATCCTGGCCGGCAAATTCGCGCC
GGGTGCCAGTATCTCGGCGAAGGTGGAAGGCGACGAGATCGTCTTCGCCTGACCTCTTCGGGGGCGGACAGGGAAAGCCC
CGCTTCGGCGGGGCTTTTTCATGGGCGTCCGGTAGGCGCGAGAGCGCTGCCGCATGCAGTGGAAGGGCCGTTTTGTCGCT
CGTGGAAAAAGAAGAAAAACTTTTTGGATCAGAGGGTTGACATGCCCGCGCGCATCCGTAAAATGGCGCGCCTCTGA

>ORF23556 (SEQ ID NO:278)
AAAAGCCCCGCCGAAGCGGGGCTTTCCCTGTCCGCCCCCGAAGAGGTCAGGCGAAGACGATCTCGTCGCCTTCCACCTTC
GCCGAGATACTGGCACCCGGCGCGAATTTGCCGGCCAGGATCAGTTGCGCCAGCGGGTTCTCGATCCAGCGCTGGATGGC
CCGCTTCAGCGGGCGTGCGCCATAGACCGGGTCGAAGCCGACGGCAATCAGCTTGTCCAGCGCCTCCTGGCTCAGTTCCA
GGCTCAGCTCGCGCTCGGCCAGGCGCTTGCGCAGGCGACCGAGCTGGATCTCGGCGATGCCGGCGATCTGCTCGCGAGCC
AGCGGCTCGAACACCACCACTTCGTCGATCCGGTTGA

Fig. 3-21

>ORF26191c (SEQ ID NO:280)
AAGGAAGGACGACCCATGCGAATAGACCGTTTGACCAGCAAGCTGCAACTGGCGCTCTCCGACGCCCAGTCCCTGGCCGT
TGGCCATGACCATCCGGCCATCGAGCCGGTGCACCTGCTTTCCGCCCTGCTCGAGCAGCAAGGCGGTTCGATCAAGCCCC
TGCTGATGCAGGTCGGCTTCGATATCGCCGCCCTGCGCAGCGGCCTCAACAAAGAACTCGACGCGCTGCCGAAGATCCAG
AGCCCGACCGGCGACGTGAACCTGTCCCAGGATCTCGCACGCCTGCTCAACCAGGCTGACCGCCTGGCCCAGCAGAAGGG
CGACCAGTTCATCTCCAGCGAGCTGGTATTGCTGGCCGCGATGGACGAGAACACCAGGCTCGGCAAGCTGCTGCTCGGCC
AGGGCGTGTCGCGCAAGGCGCTGGAGAATGCCGTGGCCAACCTGCGTGGCGGCGAAGCGGTGAACGACCCGAACGTCGAG
GAGTCGCGCCAGGCGCTGGACAAGTACACCGTCGACATGACCAAGCGCGCCGAGGAAGGCAAGCTCGACCCGGTGATCGG
TCGCGACGACGAGATCCGCCGGACCATCCAGGTCCTGCAGCGGCGGACCAAGAACAACCCGGTGCTGATCGGCGAACCCG
GCGTCGGCAAGACCGCCATCGTCGAGGGCCTGGCCCAGCGCATCATCAACGGCGAAGTGCCGGACGGCCTCAAGGACAAG
CGCCTGCTGGCCCTGGACATGGGGGGCGCTGATCGCCGGTGCCAAGTTCCGCGGCGAGTTCGAGGAACGCCTGAAGGCGGT
CCTCAACGAACTGGGCAAGCAGGAAGGCCGGGTCATCCTGTTCATCGACGAACTGCACACCATGGTCGGCGCCGGCAAGG
CGGAAGGTGCCATGGACGCCGGCAACATGCTCAAGCCGGCTCTGGCGCGCGGCGAGCTGCACTGCGTCGGTGCTACTACC
CTCGACGAGTATCGCCAGTACATCGAGAAGGATGCCGCGCTGGAGCGCCGCTTCCAGAAGGTGCTGGTGGACGAACCGAG
CGAGGAAGACACCATCGCCATCCTCCGTGGCCTCAAGGAACGCTATGAAGTGCACCACGGGGTGAGCATCACCGACGGCG
CGATCATCGCCGCGGCCAAGCTGTCGCACCGCTACATCACCGATCGGCAACTGCCGGACAAGGCCATCGACCTGATCGAC
GAGGCCGCCAGCCGCATCCGCATGGAGATCGACTCCAAGCCGGAGGAACTGGATCGTCTCGACCGTCGCCTGATCCAGCT
GAAGATCGAGCGCGAGGCGCTGAAGAAGGAAGACGACGAAGCCACCAGGAAGCGCCTGGCCAAGCTGGAGGAGGATATCG
TCAAGCTCGAGCGCGAATACGCCGACCTCGAGGAGATCTGGAAGTCCGAGAAGGCCGAGGTGCAGGGCTCGGCGCAGATC
CAGCAGAAGATCGAGCAGGCCAAGCAGGAGATGGAGGCGGCGCGGCGCAAGGGCGACCTCGAGAGCATGGCGCGCATCCA
GTACCAGACCATCCCCGGACCTGGAACGCAGCCTGCAGATGGTCGACCAGCACGGCAAGACCGAGAACCAGTTGCTGCGCA
ACAAGGTGACCGACGAGGAAATCGCCGAAGTGGTTTCCAAGTGGACCGGTATCCCGGTGTCGAAGATGCTCGAGGGCGAG
CGCGAGAAGCTGCTGCGCATGGAGCAGGAGCTGCATCGGCGAGTGATCGGCCAGGACGAGGCGGTAGTCGCCGTGTCCAA
CGCCGTGCGCCGTTCGCGCGCCGGCCTCGCCGATCCGAACCGGCCGAGCGGCTCGTTCCTCTTCCTCGGCCCGACCGGGG
TGGGCAAGACCGAGTTGTGCAAGGCGCTGGCCGAGTTCCTCTTCGATACCGAGGAGGCGCTGGTGCGGATAGATATGTCC
GAGTTCATGGAGAAACACTCGGTGGCCCGCCTGATCGGCGCGCCTCCGGGCTACGTCGGCTTCGAGGAAGGCGGCTACCT
GACCGAGGCGATCCGCCGCAAGCCCTACTCGGTGGTGCTGCTGGACGAGGTGGAGAAGGCCCATCCGGATGTATTCAACA
TTCTCCTCCAGGTGCTCGAGGACGGACGCCTGACCGACAGTCACGGGCGTACGGTGGACTTCCGCAACACCGTGGTGGTG
ATGACCTCCAACCTCGGTTCGGCGCAGATCCAGGAGCTGGCCGGCGACCGCGAGGCGCAACGTGCCGCAGTGATGGACGC
GGTCAATGCGCACTTCCGTCCGGAATTCATCAACCGGATCGACGAAGTGGTGGTGTTCGAGCCGCTGGCTCGCGAGCAGA
TCGCCGGCATCGCCGAGATCCAGCTCGGTCGCCTGCGCAAGCGCCTGGCCGAGCGCGAGCTGAGCCTGGAACTGAGCCAG
GAGGCGCTGGACAAGCTGATTGCCGTCGGCTTCGACCCGGTCTATGGCGCACGCCCGCTGAAGCGGGCCATCCAGCGCTG
GATCGAGAACCCGCTGGCGCAACTGATCCTGGCCGGCAAATTCGCGCCGGGTGCCAGTATCTCGGCGAAGGTGGAAGGCG
ACGAGATCGTCTTCGCCTGA

>ORF23751 (SEQ ID NO:282)
ACCGGGTCGAAGCCGACGGCAATCAGCTTGTCCAGCGCCTCCTGGCTCAGTTCCAGGCTCAGCTCGCGCTCGGCCAGGCG
CTTGCGCAGGCGACCGAGCTGGATCTCGGCGATGCCGGCGATCTGCTCGCGAGCCAGCGGCTCGAACACCACCACTTCGT
CGATCCGGTTGATGAATTCCGGACGGAAGTGCGCATTGACCGCGTCCATCACTGCGGCACGTTGCGCCTCGCGGTCGCCG
GCCAGCTCCTGGATCTGCGCCGAACCGAGGTTGGAGGTCATCACCACCACGGTGTTGCGGAAGTCCACCGTACGCCCGTG
A

>ORF24222 (SEQ ID NO:284)
CCCGGAGGCGCGCCGATCAGGCGGGCCACCGAGTGTTTCTCCATGAACTCGGACATATCTATCCGCACCAGCGCCTCCTC
GGTATCGAAGAGGAACTCGGCCAGCGCCTTGCACAACTCGGTCTTGCCCACCCCGGTCGGGCCGAGGAAGAGGAACGAGC
CGCTCGGCCGGTTCGGATCGGCGAGGCCGGCGCGCGAACGGCGCACGGCGTTGGACACGGCGACTACCGCCTCGTCCTGG
CCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGCTTCTCGCGCTCGCCCTCGAGCATCTTCGACACCGGGAT
ACCGGTCCACTTGGAAACCACTTCGGCGATTTCCTCGTCGGTCACCTTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGT
GCTGGTCGACCATCTGCAGGCTGCGTTCCAGGTCCGGGATGGTCTGGTACTGGATGCGCGCCATGCTCTCGAGGTCGCCC
TTGCGCCGCGCCGCCTCCATCTCCTGCTTGGCCTGCTCGATCTTCTGCTGGATCTGCGCCGAGCCCTGCACCTCGGCCTT
CTCGGACTTCCAGATCTCCTCGAGGTCGGCGTATTCGCGCTCGAGCTTGACGATATCCTCCTCCAGCTTGGCCAGGCGCT
TCCTGGTGGCTTCGTCGTCTTCCTTCTTCAGCGCCTCGCGCTCGATCTTCAGCTGGATCAGGCGACGGTCGAGACGATCC
AGTTCCTCCGGCTTGGAGTCGATCTCCATGCGGATGCGGCTGGCGGCCTCGTCGATCAGGTCGATGGCCTTGTCCGGCAG
TTGCCGATCGGTGATGTAG

Fig. 3-22

>ORF24368 (SEQ ID NO:286)
ACTCGGACATATCTATCCGCACCAGCGCCTCCTCGGTATCGAAGAGGAACTCGGCCAGCGCCTTGCACAACTCGGTCTTG
CCCACCCCGGTCGGGCCGAGGAAGAGGAACGAGCCGCTCGGCCGGTTCGGATCGGCGAGGCCGGCGCGCGAACGGCGCAC
GGCGTTGGACACGGCGACTACCGCCTCGTCCTGGCCGATCACTCGCCGATGCAGCTCCTGCTCCATGCGCAGCAGCTTCT
CGCGCTCGCCCTCGAGCATCTTCGACACCGGGATACCGGTCCACTTGGAAACCACTTCGGCGATTTCCTCGTCGGTCACC
TTGTTGCGCAGCAACTGGTTCTCGGTCTTGCCGTGCTGGTCGACCATCTGCAGGCTGCGTTCCAGGTCCGGGATGGTCTG
GTACTGGATGCGCGCCATGCTCTCGAGGTCGCCCTTGCGCCGCGCCGCCTCCATCTCCTGCTTGGCCTGCTCGATCTTCT
GCTGGATCTGCGCCGAGCCCTGCACCTCGGCCTTCTCGGACTTCCAGATCTCCTCGAGGTCGGCGTATTCGCGCTCGAGC
TTGA

>ORF24888c (SEQ ID NO:288)
AGAAGGAAGACGACGAAGCCACCAGGAAGCGCCTGGCCAAGCTGGAGGAGGATATCGTCAAGCTCGAGCGCGAATACGCC
GACCTCGAGGAGATCTGGAAGTCCGAGAAGGCCGAGGTGCAGGGCTCGGCGCAGATCCAGCAGAAGATCGAGCAGGCCAA
GCAGGAGATGGAGGCGGCGCGGCGCAAGGGCGACCTCGAGAGCATGGCGCGCATCCAGTACCAGACCATCCCGGACCTGG
AACGCAGCCTGCAGATGGTCGACCAGCACGGCAAGACCGAGAACCAGTTGCTGCGCAACAAGGTGA

>ORF25398c (SEQ ID NO:290)
AGGCGGTCCTCAACGAACTGGGCAAGCAGGAAGGCCGGGTCATCCTGTTCATCGACGAACTGCACACCATGGTCGGCGCC
GGCAAGGCGGAAGGTGCCATGGACGCCGGCAACATGCTCAAGCCGGCTCTGGCGCGCGGCGAGCTGCACTGCGTCGGTGC
TACTACCCTCGACGAGTATCGCCAGTACATCGAGAAGGATGCCGCGCTGGAGCGCCGCTTCCAGAAGGTGCTGGTGGACG
AACCGAGCGAGGAAGACACCATCGCCATCCTCCGTGGCCTCAAGGAACGCTATGAAGTGCACCACGGGGTGA

>ORF25892c (SEQ ID NO:292)
CCGCCTGGCCCAGCAGAAGGGCGACCAGTTCATCTCCAGCGAGCTGGTATTGCTGGCCGCGATGGACGAGAACACCAGGC
TCGGCAAGCTGCTGCTCGGCCAGGGCGTGTCGCGCAAGGCGCTGGAGAATGCCGTGGCCAACCTGCGTGGCGGCGAAGCG
GTGAACGACCCGAACGTCGAGGAGTCGCGCGCCAGGCGCTGGACAAGTACACCGTCGACATGACCAAGCGCGCCGAGGAAGG
CAAGCTCGACCCGGTGATCGGTCGCGACGACGAGATCCGCCGGACCATCCAGGTCCTGCAGCGGCGGACCAAGAACAACC
CGGTGCTGATCGGCGAACCCGGCGTCGGCAAGACCGCCATCGTCGAGGGCCTGGCCCAGCGCATCATCAACGGCGAAGTG
CCGGACGGCCTCAAGGACAAGCGCCTGCTGGCCCTGGACATGGGGGCGCTGATCGCCGGTGCCAAGTTCCGCGGCGAGTT
CGAGGAACGCCTGAAGGCGGTCCTCAACGAACTGGGCAAGCAGGAAGGCCGGGTCATCCTGTTCATCGACGAACTGCACA
CCATGGTCGGCGCCGGCAAGGCGGAAGGTGCCATGGACGCCGGCAACATGCTCAAGCCGGCTCTGGCGCGCGGCGAGCTG
CACTGCGTCGGTGCTACTACCCTCGACGAGTATCGCCAGTACATCGAGAAGGATGCCGCGCTGGAGCGCCGCTTCCAGAA
GGTGCTGGTGGACGAACCGAGCGAGGAAGACACCATCGCCATCCTCCGTGGCCTCAAGGAACGCTATGA

>ORF25110 (SEQ ID NO:294)
CGTTCCTTGAGGCCACGGAGGATGGCGATGGTGTCTTCCTCGCTCGGTTCGTCCACCAGCACCTTCTGGAAGCGGCGCTC
CAGCGCGGCATCCTTCTCGATGTACTGGCGATACTCGTCGAGGGTAGTAGCACCGACGCAGTGCAGCTCGCCGCGCGCCA
GAGCCGGCTTGAGCATGTTGCCGGCGTCCATGGCACCTTCCGCCTTGCCGGCGCCGACCATGGTGTGCAGTTCGTCGATG
AACAGGATGACCCGGCCTTCCTGCTTGCCCAGTTCGTTGAGGACCGCCTTCAGGCGTTCCTCGAACTCGCCGCGGAACTT
GGCACCGGCGATCAGCGCCCCCATGTCCAGGGCCAGCAGGCGCTTGTCCTTGAGGCCGTCCGGCACTTCGCCGTTGATGA
TGCGCTGGGCCAGGCCCTCGACGATGGCGGTCTTGCCGACGCCGGGTTCGCCGATCAGCACCGGGTTGTTCTTGGTCCGC
CGCTGCAGGACCTGGATGGTCCGGCGGATCTCGTCGTCGCGACCGATCACCGGGTCGAGCTTGCCTTCCTCGGCGCGCTT
GGTCATGTCGACGGTGTACTTGTCCAGCGCCTGGCGCGACTCCTCGACGTTCGGGTCGTTCACCGCTTCGCCGCCACGCA
GGTTGGCCACGGCATTCTCCAGCGCCTTGCGCGACACGCCCTGGCCGAGCAGCAGCTTGCCGAGCCTGGTGTTCTCGTCC
ATCGCGGCCAGCAATACCAGCTCGCTGGAGATGAACTGGTCGCCCTTCTGCTGGGCCAGGCGGTCAGCCTGGTTGAGCAG
GCGTGCGAGATCCTGGGACAGGTTCACGTCGCCGGTCGGGCTCTGGATCTTCGGCAGCGCGTCGAGTTCTTTGTTGAGGC
CGCTGCGCAGGGCGGCGATATCGAAGCCGACCTGCATCAGCAGGGGCTTGATCGAACCGCCTTGCTGCTCGAGCAGGGCG
GAAAGCAGGTGCACCGGCTCGATGGCCGGATGGTCATGGCCAACGGCCAGGGACTGGGCGTCGGAGAGCGCCAGTTGCAG
CTTGCTGGTCAAACGGTCTATTCGCATGGGTCGTCCTTCCTTCTATAGAGCGGGCCGGAACGATGGGTGTCCCTGA

>ORF25510 (SEQ ID NO:296)
TGCGCTGGGCCAGGCCCTCGACGATGGCGGTCTTGCCGACGCCGGGTTCGCCGATCAGCACCGGGTTGTTCTTGGTCCGC
CGCTGCAGGACCTGGATGGTCCGGCGGATCTCGTCGTCGCGACCGATCACCGGGTCGAGCTTGCCTTCCTCGGCGCGCTT
GGTCATGTCGACGGTGTACTTGTCCAGCGCCTGGCGCGACTCCTCGACGTTCGGGTCGTTCACCGCTTCGCCGCCACGCA
GGTTGGCCACGGCATTCTCCAGCGCCTTGCGCGACACGCCCTGGCCGAGCAGCAGCTTGCCGAGCCTGGTGTTCTCGTCC
ATCGCGGCCAGCAATACCAGCTCGCTGGAGATGA

Fig. 3-23

>ORF26762c (SEQ ID NO:298)
CCGCCGACTGCCTGCCGGCGTTGTTCTGCGACCGCTCGGGCACCCGGGTGGCCGCGGCCCATGCCGGCTGGCGCGGGCTG
GCGGCGGGCGTGCTGGAGGCGACGGTGGACAGCCTGGGCGTGCCCGGCGACGAACTGCTGGTCTGGCTGGGGCCGGCGAT
CGGCCCGCAGGCCTTCGAGGTCGGCGGCGAGGTCCGCGATGCATTCGTCGCTGCGCACGCCGAGGCGCGCTCGGCTTTCG
TACCTAGCGCCAATCCGGGCCGCTTCATGGCCGACATCTACCGACTCGCGCGGATCCGCCTGGGCGCCCATGGCGTCACC
GCCGTGCATGGCGGCGGCTTCTGCACCTTCAGCGATACCGCGCGCTTCTATTCCTACCGCCGCTCGTCGCGTACCGGCCG
TTTTGCCAGCCTGGTCTGGCTCCAGGACTAGGCCCGCGCAGGTTATCCGGCGGCAACTGACCGATGTCACGGTCCGGTCG
CTTGAACCGCGGAAAATCGCCCTTATCTACTGA

>ORF26257 (SEQ ID NO:300)
ATAAGGGCGATTTTCCGCGGTTCAAGCGACCGGACCGTGACATCGGTCAGTTGCCGCCGGATAACCTGCGCGGGCCTAGT
CCTGGAGCCAGACCAGGCTGGCAAAACGGCCGGTACGCGACGAGCGGCGGTAGGAATAGAAGCGCGCGGTATCGCTGAAG
GTGCAGAAGCCGCCGCCATGCACGGCGGTGACGCCATGGGCGCCCAGGCGGATCCGCGCGAGTCGGTAGATGTCGGCCAT
GAAGCGGCCCGGATTGGCGCTAGGTACGAAAGCCGAGCGCGCCTCGGCGTGCGCAGCGACGAATGCATCGCGGACCTCGC
CGCCGACCTCGAAGGCCTGCGGGCCGATCGCCGGCCCCAGCCAGACCAGCAGTTCGTCGCCGGGCACGCCCAGGCTGTCC
ACCGTCGCCTCCAGCACGCCCGCCGCCAGCCCGCGCCAGCCGGCATGGGCCGCGGCCACCCGGGTGCCCGAGCGGTCGCA
GAACAACGCCGGCAGGCAGTCGGCGGTCATGATCGTACAGGCGACGCCCGGCATCGCGCTCCAGCTGGCGTCGGCCCTGA
GCACCGGTTCGGGTCGGCCTCCACCACGTCACTCCGTGCACCTATTCCAACCAGCTCGGCCGGCATTCCAGACGCTCGGT
CAGGCGTCGGCGGTTTTATTCCACGGCGCGCGGATCGTCGTAGACGTGGGCGCCAAGGTTCAGACTGTCGAAGGGTGCCT
GGCTGA

>ORF26844c (SEQ ID NO:302)
CGTGGTGGAGGCCGACCCGAACCGGTGCTCAGGGCCGACGCCAGCTGGAGCGCGATGCCGGGCGTCGCCTGTACGATCAT
GACCGCCGACTGCCTGCCGGCGTTGTTCTGCGACCGCTCGGGCACCCGGGTGGCCGCGGCCCATGCCGGCTGGCGCGGGC
TGGCGGCGGGCGTGCTGGAGGCGACGGTGGACAGCCTGGGCGTGCCCGGCGACGAACTGCTGGTCTGGCTGGGGCCGGCG
ATCGGCCCGCAGGCCTTCGAGGTCGGCGGCGAGGTCCGCGATGCATTCGTCGCTGCGCACGCCGAGGCGCGCTCGGCTTT
CGTACCTAGCGCCAATCCGGGCCGCTTCATGGCCGACATCTACCGACTCGCGCGGATCCGCCTGGGCGCCCATGGCGTCA
CCGCCGTGCATGGCGGCGGCTTCTGCACCTTCAGCGATACCGCGCGCTTCTATTCCTACCGCCGCTCGTCGCGTACCGGC
CGTTTTGCCAGCCTGGTCTGGCTCCAGGACTAG

>ORF26486 (SEQ ID NO:304)
ATGTCGGCCATGAAGCGGCCCGGATTGGCGCTAGGTACGAAAGCCGAGCGCGCCTCGGCGTGCGCAGCGACGAATGCATC
GCGGACCTCGCCGCCGACCTCGAAGGCCTGCGGGCCGATCGCCGGCCCCAGCCAGACCAGCAGTTCGTCGCCGGGCACGC
CCAGGCTGTCCACCGTCGCCTCCAGCACGCCCGCCGCCAGCCCGCGCCAGCCGGCATGGGCCGCGGCCACCCGGGTGCCC
GAGCGGTCGCAGAACAACGCCGGCAGGCAGTCGGCGGTCATGATCGTACAGGCGACGCCCGGCATCGCGCTCCAGCTGGC
GTCGGCCCTGAGCACCGGTTCGGGTCGGCCTCCACCACGTCACTCCGTGCACCTATTCCAACCAGCTCGGCCGGCATTCC
AGACGCTCGGTCAGGCGTCGGCGGTTTTATTCCACGGCGCGCGGATCGTCGTAGACGTGGGCGCCAAGGTTCAGACTGTC
GAAGGGTGCCTGGCTGACCCCGCCACTGCGCGTGGTCACGCAGGCCCGCACACGGGCCGGCGCCGGCCAGTCGGGGGTCA
GCCAGGCGTTCAACCGACGAACGCCTCGCGATCCTGGCGCAACAGGCTGAGCAGCCAGAGGAATTCTTCCGGCAGCGGCG
ATTCCCACTTCATGCGCACGCCGGTGGCCGGGTGA

>ORF26857c (SEQ ID NO:306)
GTGCACGGAGTGACGTGGTGGAGGCCGACCCGAACCGGTGCTCAGGGCCGACGCCAGCTGGAGCGCGATGCCGGGCGTCG
CCTGTACGATCATGACCGCCGACTGCCTGCCGGCGTTGTTCTGCGACCGCTCGGGCACCCGGGTGGCCGCGGCCCATGCC
GGCTGGCGCGGGCTGGCGGCGGGCGTGCTGGAGGCGACGGTGGACAGCCTGGGCGTGCCCGGCGACGAACTGCTGGTCTG
GCTGGGGCCGGCGATCGGCCCGCAGGCCTTCGAGGTCGGCGGCGAGGTCCGCGATGCATTCGTCGCTGCGCACGCCGAGG
CGCGCTCGGCTTTCGTACCTAG

>ORF27314c (SEQ ID NO:308)
AGTGGGAATCGCCGCTGCCGGAAGAATTCCTCTGGCTGCTCAGCCTGTTGCGCCAGGATCGCGAGGCGTTCGTCGGTTGA
ACGCCTGGCTGACCCCCGACTGGCCGGCGCCGGCCCGTGTGCGGGCCTGCGTGACCACGCGCAGTGGCGGGGTCAGCCAG
GCACCCTTCGACAGTCTGAACCTTGGCGCCCACGTCTACGACGATCCGCGCGCCGTGGAATAAAACCGCCGACGCCTGAC
CGAGCGTCTGGAATGCCGGCCGAGCTGGTTGGAATAGGTGCACGGAGTGACGTGGTGGAGGCCGACCCGAACCGGTGCTC
AGGGCCGACGCCAGCTGGAGCGCGATGCCGGGCGTCGCCTGTACGATCATGA

Fig. 3-24

>ORF27730c (SEQ ID NO:310)
CAAGCCCGCCGGCCTGGTGGTCCATCCGGCTGCCGGCCATCAGGACGGCACCCTGCTGAATGCCTTGCTCTACCATGTCC
CGGACATCGCCAATGTGCCGCGCGCCGGGATCGTCCACCGCCTGGACAAGGACACGACCGGCCTGATGGTAGTGGCCAAG
ACGCTGGAGGCCCACACCAAGCTGGTGGCGCAACTGCAGGCACGGTCGGTCAGCCGCATCTACGAGGCGATCGTGATCGG
CGTGATCACCTCCGGCGGCACCATCGATGCGCCGATCGGACGGCATGGCGTGCAGCGGCAGAAGATGGCGGTGGTCGACG
CCGGCAAGGTGGCGGTCAGCCATTACCGCGTGCTGGAACGCTTCCGTGCGCACACCCATACCCGGGTCAAGCTGGAGACC
GGGCGTACCCACCAGATCCGCGTGCACATGAGCCATATTGGCTATCCCCTGGTCGGCGATCCGGTCTACGGTGGGCGCTT
CAGGATTCCCCCGGTGGCCAGCCAGACCCTGGTCCAGACTCTTCGCGAATTCCCCCGGCAGGCGCTGCACGCGCGCTTCC
TCGAACTGGATCACCCGGCCACCGGCGTGCGCATGAAGTGGGAATCGCCGCTGCCGGAAGAATTCCTCTGGCTGCTCAGC
CTGTTGCGCCAGGATCGCGAGGCGTTCGTCGGTTGAACGCCTGGCTGACCCCCGACTGGCCGGCGCCGGCCCGTGTGCGG
GCCTGCGTGACCACGCGCAGTGGCGGGGTCAGCCAGGCACCCTTCGACAGTCTGAACCTTGGCGCCCACGTCTACGACGA
TCCGCGCGCCGTGGAATAA

>ORF26983 (SEQ ID NO:312)
CCCCGCCACTGCGCGTGGTCACGCAGGCCCGCACACGGGCCGGCGCCGGCCAGTCGGGGGTCAGCCAGGCGTTCAACCGA
CGAACGCCTCGCGATCCTGGCGCAACAGGCTGAGCAGCCAGAGGAATTCTTCCGGCAGCGGCGATTCCCACTTCATGCGC
ACGCCGGTGGCCGGGTGATCCAGTTCGAGGAAGCGCGCGTGCAGCGCCTGCCGGGGGGAATTCGCGAAGAGTCTGGACCAG
GGTCTGGCTGGCCACCGGGGGAATCCTGAAGCGCCCACCGTAGACCGGATCGCCGACCAGGGGATAGCCAATATGGCTCA
TGTGCACGCGGATCTGGTGGGTACGCCCGGTCTCCAGCTTGACCCGGGTATGGGTGTGCGCACGGAAGCGTTCCAGCACG
CGGTAATGGCTGACCGCCACCTTGCCGGCGTCGACCACCGCCATCTTCTGCCGCTGCACGCCATGCCGTCCGATCGGCGC
ATCGATGGTGCCGCCGGAGGTGATCACGCCGATCACGATCGCCTCGTAGATGCGGCTGACCGACCGTGCCTGCAGTTGCG
CCACCAGCTTGGTGTGGGCCTCCAGCGTCTTGGCCACTACCATCAGGCCGGTCGTGTCCTTGTCCAGGCGGTGGACGATC
CCGGCGCGCGGCACATTGGCGATGTCCGGGACATGGTAGAGCAAGGCATTCAGCAGGGTGCCGTCCTGATGGCCGGCAGC
CGGATGGACCACCAGGCCGGCGGGCTTGTCAATCACCAGGATGTGCTCGTCCTCGTAGACGATTTCCAGCTCGATGTCCT
GTGCGAGCCACTCGCCCTGGGCTTCCTGCTCGGCCTCCAGGACCAGTTGCGCGCCGCTGTGGACGATGTCGCGCGGGCGC
AGCACGGCGCCGTCGACGGTCAGGCGACCGTCCTTGATCCAGCCGGCCAGACGGGAGCGGGAGTGTTCGGGAAAAAGCTG
GGCGGCGATCTGGTCGAGACGCTGGCCACCCAGCTCGAACGGCACCTCGGCCGCGCGTTGAATCATATCGGACATGAGTA
G

>ORF28068c (SEQ ID NO:314)
CCACAGCGCGTAGCCGATTCCAAAAGCCGCGCTGAGCATCGTCTCCTACTCATGTCCGATATGATTCAACGCGCGGCCGA
GGTGCCGTTCGAGCTGGGTGGCCAGCGTCTCGACCAGATCGCCGCCCAGCTTTTTCCCGAACACTCCCGCTCCCGTCTGG
CCGGCTGGATCAAGGACGGTCGCCTGACCGTCGACGGCGCCGTGCTGCGCCCCGCGCGACATCGTCCACAGCGGCGCGCAA
CTGGTCCTGGAGGCCGAGCAGGAAGCCCAGGGCGAGTGGCTCGCACAGGACATCGAGCTGGAAATCGTCTACGAGGACGA
GCACATCCTGGTGATTGACAAGCCCGCCGGCCTGGTGGTCCATCCGGCTGCCGGCCATCAGGACGGCACCCTGCTGAATG
CCTTGCTCTACCATGTCCCGGACATCGCCAATGTGCCGCGCGCCGGGATCGTCCACCGCCTGGACAAGGACACGACCGGC
CTGATGGTAGTGGCCAAGACGCTGGAGGCCCACACCAAGCTGGTGGCGCAACTGCAGGCACGGTCGGTCAGCCGCATCTA
CGAGGCGATCGTGATCGGCGTGATCACCTCCGGCGGCACCATCGATGCGCCGATCGGACGGCATGGCGTGCAGCGGCAGA
AGATGGCGGTGGTCGACGCCGGCAAGGTGGCGGTCAGCCATTACCGCGTGCTGGAACGCTTCCGTGCGCACACCCATACC
CGGGTCAAGCTGGAGACCGGGCGTACCCACCAGATCCGCGTGCACATGAGCCATATTGGCTATCCCCTGGTCGGCGATCC
GGTCTACGGTGGGCGCTTCAGGATTCCCCCGGTGGCCAGCCAGACCCTGGTCCAGACTCTTCGCGAATTCCCCCGGCAGG
CGCTGCACGCGCGCTTCCTCGAACTGGATCACCCGGCCACCGGCGTGCGCATGAAGTGGGAATCGCCGCTGCCGGAAGAA
TTCCTCTGGCTGCTCAGCCTGTTGCGCCAGGATCGCGAGGCGTTCGTCGGTTGA

>ORF27522 (SEQ ID NO:316)
CCGACCGTGCCTGCAGTTGCGCCACCAGCTTGGTGTGGGCCTCCAGCGTCTTGGCCACTACCATCAGGCCGGTCGTGTCC
TTGTCCAGGCGGTGGACGATCCCGGCGCGCGGCACATTGGCGATGTCCGGGACATGGTAGAGCAAGGCATTCAGCAGGGT
GCCGTCCTGATGGCCGGCAGCCGGATGGACCACCAGGCCGGCGGGCTTGTCAATCACCAGGATGTGCTCGTCCTCGTAGA
CGATTTCCAGCTCGATGTCCTGTGCGAGCCACTCGCCCTGGGCTTCCTGCTCGGCCTCCAGGACCAGTTGCGCGCCGCTG
TGGACGATGTCGCGCGGGCGCAGCACGGCGCCGTCGACGGTCAGGCGACCGTCCTTGA

Fig. 3-25

>ORF28033c (SEQ ID NO:318)
GCATCGTCTCCTACTCATGTCCGATATGATTCAACGCGCGGCCGAGGTGCCGTTCGAGCTGGGTGGCCAGCGTCTCGACC
AGATCGCCGCCCAGCTTTTTCCCGAACACTCCCGCTCCCGTCTGGCCGGCTGGATCAAGGACGGTCGCCTGACCGTCGAC
GGCGCCGTGCTGCGCCCGCGCGACATCGTCCACAGCGGCGCGCAACTGGTCCTGGAGGCCGAGCAGGAAGCCCAGGGCGA
GTGGCTCGCACAGGACATCGAGCTGGAAATCGTCTACGAGGACGAGCACATCCTGGTGATTGA

>ORF29701c (SEQ ID NO:320)
TCTTCCAGTTCGCTGGAGATCAGCAGGACCAGTACCAGGCCGATGGTCAGGCGGTACAGGTGGTACAGACGGAGGATGCG
TTGCCCCTGCTCCTCGCTCAGCCGTAGCCGTTCAGCGCGCACGGTCGCCCTGGTCCTGGCGCAGGTGCGCCTGGCTGCAA
TACCAGCGTTGTTCGTGGGCGAGGGCGTTGGCCTGCGGCACGTGGACGCCGCAATGGGCGCAGCGGACCATCGGCGATGC
GCTCGGCTCGTCCTGCGGACGTTGCTGCTGGCGCGGAGTGGGACGGGTAAAGCGACGCCAGAGCCAGAACGCGATGGCGA
TCAGGGCGATCCAGAACAGGAGGCGGAAAAGGCCCATGGTGATCTCGGAGGCTGGAGAAAGCTGCAGTTTAGCCAAGCCG
CCGGCTCGATCCCAGACGGGAAGGTCCAGGCTGTGCGGCGTTTGGCGCTGGGAGAGGCATGGCGGCGGGCAAAAAGAAGG
GAGGCCTGCGCCTCCCTTCGGTGTTTCGTGCGATCAGTCGAAGAGACCGAAGGTCATGTAGCTCCACCAGGAGCGACCGG
AGTCCTCGTCGTCATCGCTCTCCGGCTTCTCGTCGTCGGCGCTGTGATCCTGGTTTTCCGGCTTCAGTTCGGCGGGGATC
TCCCGCTCGGCATCCTCGTACTGCTTGATCACGTCCTTGGCGGCCTGGGTTTCCATGTGCGGCGGCGGCTCGCCGCCTTC
GATCAGGCCCAGGGTGGCCTTGGCCAGCCAGGAGCGGGTGTCGGCCTCGCTTTCGCGGGCGACGAACTCGCCATCCTTGA
GGCTGGCGTTATCCGGATAGTTCAGCTTGAGGGTTTCCAGGCTGGTGCTGGCCAGGTCGTCGAGACCCAGGCGACGGTAG
GCTTCGACCATGATCGCCAGGCCATCGCCGACGGCCGGGGTTTCCTGGAAGTTCTCCACCACGTAGCGACCGCGGTTGGC
GGCGGCGACATAGGCCTGGCGCTTCAGGTAGTAGTGGCCGACGTGCACTTCGTAGGCCGCCAGCAGGTTGCGCAGGTACA
CCATGCGCGCCTTGGCGTCCGGGGCGTAGCGGCTGTTGGGGAAGCGGCTGGTGAGCTGGGCGAACTCGTTGAAGGAGTCG
CGGGCGGCGCCCGGGTCGCGCTTGGTCATGTCCAGCGGCAGGAAGCGCGCCAGCAGGCCGCGGTCCTGGTCGAAGGAGGA
CAGGCCTTTGAGGTAGTAGGCGTAGTCGACGTTGGGGTGCTGCGGATGCAGGCGGATGAAGCGTTCGGCGGCGGCGCGGG
CGGCTTCGGGCTCCATGTTCTTGTAGTTGGCGTAGATCAGCTCGAGCTGGGCCTGCTCGGCGTAGCGGCCGAAGGGATAG
CGCGATTCGAGGGCTTTCAGCTTGGTGACGGCGCTGTTGTAGCTCTTGTTGTTGAGGTCGTCCTGCGCCTGCTGGTACAG
CTGGCTCTCGCTCAGGTTCTCGTCGACAGTCTCCTTGTTCGAGGAGCAGGCTGCGGTGAGGGCGAGGATGGCGATCAGCA
GCAGGTGTTTCACTTGCATGGCGGCTTGCGTCCCTGGGACGGTCGGCTTGGCCTCAACCGTCTGTTATGA

>ORF28118 (SEQ ID NO:322)
CAGACGGTTGAGGCCAAGCCGACCGTCCCAGGGACGCAAGCCGCCATGCAAGTGAAACACCTGCTGCTGATCGCCATCCT
CGCCCTCACCGCAGCCTGCTCCTCGAACAAGGAGACTGTCGACGAGAACCTGAGCGAGAGCCAGCTGTACCAGCAGGCGC
AGGACGACCTCAACAACAAGAGCTACAACAGCGCCGTCACCAAGCTGAAAGCCCTCGAATCGCGCTATCCCTTCGGCCGC
TACGCCGAGCAGGCCCAGCTCGAGCTGATCTACGCCAACTACAAGAACATGGAGCCCGAAGCCGCCCGCGCCGCCGCCGA
ACGCTTCATCCGCCTGCATCCGCAGCACCCCAACGTCGACTACGCCTACTACCTCAAAGGCCTGTCCTCCTTCGACCAGG
ACCGCGGCCTGCTGGCGCGCGCTTCCTGCCGCTGGACATGACCAAGCGCGACCCGGGCGCCGCCCGCGACTCCTTCAACGAG
TTCGCCCAGCTCACCAGCCGCTTCCCCAACAGCCGCTACGCCCCGGACGCCAAGGCGCGCATGGTGTACCTGCGCAACCT
GCTGGCGGCCTACGAAGTGCACGTCGGCCACTACTACCTGAAGCGCCAGGCCTATGTCGCCGCCGCCAACCGCGGTCGCT
ACGTGGTGGAGAACTTCCAGGAAACCCCGGCCGTCGGCGATGGCCTGGCGATCATGGTCGAAGCCTACCGTCGCCTGGGT
CTCGACGACCTGGCCAGCACCAGCCTGGAAACCCTCAAGCTGAACTATCCGGATAACGCCAGCCTCAAGGATGGCGAGTT
CGTCGCCCGCGAAAGCGAGGCCGACACCCGCTCCTGGCTGGCCAAGGCCACCCTGGGCCTGATCGAAGGCGGCGAGCCGC
CGCCGCACATGGAAACCCAGGCCGCCAAGGACGTGATCAAGCAGTACGAGGATGCCGAGCGGGAGATCCCCGCCGAACTG
AAGCCGGAAAACCAGGATCACAGCGCCGACGACGAGAAGCCGGAGAGCGATGACGACGAGGACTCCGGTCGCTCCTGGTG
GAGCTACATGACCTTCGGTCTCTTCGACTGA

>ORF28129 (SEQ ID NO:324)
GGCCAAGCCGACCGTCCCAGGGACGCAAGCCGCCATGCAAGTGAAACACCTGCTGCTGATCGCCATCCTCGCCCTCACCG
CAGCCTGCTCCTCGAACAAGGAGACTGTCGACGAGAACCTGAGCGAGAGCCAGCTGTACCAGCAGGCGCAGGACGACCTC
AACAACAAGAGCTACAACAGCGCCGTCACCAAGCTGAAAGCCCTCGAATCGCGCTATCCCTTCGGCCGCTACGCCGAGCA
GGCCCAGCTCGAGCTGATCTACGCCAACTACAAGAACATGGAGCCCGAAGCCGCCCGCGCCGCCGCCGAACGCTTCATCC
GCCTGCATCCGCAGCACCCCAACGTCGACTACGCCTACTACCTCAAAGGCCTGTCCTCCTTCGACCAGGACCGCGGCCTG
CTGGCGCGCTTCCTGCCGCTGGACATGACCAAGCGCGACCCGGGCGCCGCCCGCGACTCCTTCAACGAGTTCGCCCAGCT
CACCAGCCGCTTCCCCAACAGCCGCTACGCCCCGGACGCCAAGGCGCGCATGGTGTACCTGCGCAACCTGCTGGCGGCCT
ACGAAGTGCACGTCGGCCACTACTACCTGAAGCGCCAGGCCTATGTCGCCGCCGCCAACCGCGGTCGCTACGTGGTGGAG
AACTTCCAGGAAACCCCGGCCGTCGGCGATGGCCTGGCGATCATGGTCGAAGCCTACCGTCGCCTGGGTCTCGACGACCT
GGCCAGCACCAGCCTGGAAACCCTCAAGCTGAACTATCCGGATAA

Fig. 3-26

>ORF29709c (SEQ ID NO:326)
GGACCTGATCTTCCAGTTCGCTGGAGATCAGCAGGACCAGTACCAGGCCGATGGTCAGGCGGTACAGGTGGTACAGACGG
AGGATGCGTTGCCCCTGCTCCTCGCTCAGCCGTAGCCGTTCAGCGCGCACGGTCGCCCTGGTCCTGGCGCAGGTGCGCCT
GGCTGCAATACCAGCGTTGTTCGTGGGCGAGGGCGTTGGCCTGCGGCACGTGGACGCCGCAATGGGCGCAGCGGACCATC
GGCGATGCGCTCGGCTCGTCCTGCGGACGTTGCTGCTGGCGCGGAGTGGGACGGGTAAAGCGACGCCAGAGCCAGAACGC
GATGGCGATCAGGGCGATCCAGAACAGGAGGCGGAAAAGGCCCATGGTGATCTCGGAGGCTGGAGAAAGCTGCAGTTTAG
CCAAGCCGCCGGCTCGATCCCAGACGGGAAGGTCCAGGCTGTGCGGCGTTTGGCGCTGGGAGAGGCATGGCGGCGGGCAA
AAAGAAGGGAGGCCTGCGCCTCCCTTCGGTGTTTCGTGCGATCAGTCGAAGAGACCGAAGGTCATGTAG

>ORF29189 (SEQ ID NO:328)
TCGCACGAAACACCGAAGGGAGGCGCAGGCCTCCCTTCTTTTTTGCCCGCCGCCATGCCTCTCCCAGCGCCAAACGCCGCA
CAGCCTGGACCTTCCCGTCTGGGATCGAGCCGGCGGCTTGGCTAAACTGCAGCTTTCTCCAGCCTCCGAGATCACCATGG
GCCTTTTCCGCCTCCTGTTCTGGATCGCCCTGATCGCCATCGCGTTCTGGCTCTGGCGTCGCTTTACCCGTCCCACTCCG
CGCCAGCAGCAACGTCCGCAGGACGAGCCGAGCGCATCGCCGATGGTCCGCTGCGCCCATTGCGGCGTCCACGTGCCGCA
GGCCAACGCCCTCGCCCACGAACAACGCTGGTATTGCAGCCAGGCGCACCTGCGCCAGGACCAGGGCGACCGTGCGCGCT
GA

>ORF29382 (SEQ ID NO:330)
TCGCCATCGCGTTCTGGCTCTGGCGTCGCTTTACCCGTCCCACTCCGCGCCAGCAGCAACGTCCGCAGGACGAGCCGAGC
GCATCGCCGATGGTCCGCTGCGCCCATTGCGGCGTCCACGTGCCGCAGGCCAACGCCCTCGCCCACGAACAACGCTGGTA
TTGCAGCCAGGCGCACCTGCGCCAGGACCAGGGCGACCGTGCGCGCTGAACGGCTACGGCTGAGCGAGGAGCAGGGGCAA
CGCATCCTCCGTCTGTACCACCTGTACCGCCTGACCATCGGCCTGGTACTGGTCCTGCTGATCTCCAGCGAACTGGAAGA
TCAGGTCCTCAAGCTCGTCCACCCTGAACTGTTCCATGTCGGCAGTTGGTGCTACCTGGTCTTCAACATCCTGGTCGCGC
TGTTCCTGCCGCCGTCGCGGCAATTGCTGCCGATCTTCATCCTCGCGCTCACCGACGTGCTGATGCTTTGCGGCCTGTTC
TACGCAGGTGGCGGCGTACCCAGCGGCATCGGCAGCCTGCTGGTGGTGGCGGTGGCCATTGCCAACATCCTGCTGCGCGG
GCGCATCGGCCTGGTCATCGCGGCGGCGGCCAGCCTCGGCCTGCTCTACCTGACCTTCTTCCTCAGCCTGAGCAGTCCGG
ACGCCACCAACCACTACGTCCAGGCCGGCGGCCTCGGCACCCTGTGCTTCGCCGCCGCGCTGGTGATCCAGGCTCTGGTG
CGGCGCCAGGAGCAGACCGAAACGCTGGCCGAAGAACGCGCCGAGACGGTCGCCAACCTGGAGGAACTCAACGCATTGAT
CCTGCAGCGCATGCGCACCGGCATCCTCGTGGTCGATAGCCGTCAGGCCATCCTCCTCGCCAACCAGGCCGCCCTCGGCC
TGCTCAGGCAGGACGACGTGCAGGGCGCCAGCCTCGGCCGCCACAGCCCGATGCTGATGCACTGCATGAAGCAATGGCGC
CTGAATCCCAGCCTCCGTCCGCCGACGCTCAAGGTGGTGCCGGATGGCCCGACGGTGCAACCCAGCTTTATCAGCCTCAA
CCGCGAAGACGACCAGCACGTGCTGATCTTCCTCGAAGACATTTCGCAGATCGCCCAGCAGGCGCAGCAGATGAAGCTGG
CCGGTCTTGGCCGCCTGACCGCCGGCATCGCCCATGAGATCCGCAACCCGCTGGGCGCGATCAGCCACGCCGCCCAACTG
CTGCAGGAGTCAGAGGAACTGGATGCCCCGGACCGACGCCTGACGCAGATCATCCAGGACCAGTCGAAGCGGATGAACCT
GGTCATCGAGAACGTCCTGCAGCTCTCCCGTCGCCGCCAGGCCGAACCGCAGCAGCTCGACCTGAAGGAGTGGCTTCAGC
GGTTCGTCGACGAATACCCCGGCAGGCTGCGCAACGACAGCCAACTGCACCTGCAGCTCGGTGCCGGCGACATCCAGACC
CGCATGGACCCACACCAGTTGAACCAGGTGCTGAGCAACCTGGTGCAGAACGGTCTTCGCTACAGCGCCCAGGCGCACGG
GCGCGGCCAGGTCTGGCTGAGCCTCGCGCGCGACCCGGAGAGCGACCTGCCGGTGCTGGAAGTCATCGACGACGGTCCCG
GCGTACCGGCGGACAAACTGAACAACCTGTTCGAACCCTTCTTTACTACAGAAAGCAAAGGCACCGGCCTGGGCCTCTAT
CTCTCCCGCGAACTCTGCGAGAGCAACCAGGCACGGATCGACTACCGCAATCGCGAGGAAGGCGGCGGCTGCTTCCGCAT
CACCTTCGCCCACCCGCGCAAACTCAGCTGA

Fig. 3-27

>ORF30590c (SEQ ID NO:332)
CTCCTGCAGCAGTTGGGCGGCGTGGCTGATCGCGCCCAGCGGGTTGCGGATCTCATGGGCGATGCCGGCGGTCAGGCGGC
CAAGACCGGCCAGCTTCATCTGCTGCGCCTGCTGGGCGATCTGCGAAATGTCTTCGAGGAAGATCAGCACGTGCTGGTCG
TCTTCGCGGTTGAGGCTGATAAAGCTGGGTTGCACCGTCGGGCCATCCGGCACCACCTTGAGCGTCGGCGGACGGAGGCT
GGGATTCAGGCGCCATTGCTTCATGCAGTGCATCAGCATCGGGCTGTGGCGGCCGAGGCTGGCGCCCTGCACGTCGTCCT
GCCTGAGCAGGCCGAGGGCGGCCTGGTTGGCGAGGAGGATGGCCTGACGGCTATCGACCACGAGGATGCCGGTGCGCATG
CGCTGCAGGATCAATGCGTTGAGTTCCTCCAGGTTGGCGACCGTCTCGGCGCGTTCTTCGGCCAGCGTTTCGGTCTGCTC
CTGGCGCCGCACCAGAGCCTGGATCACCAGCGCGGCGGCGAAGCACAGGGTGCCGAGGCCGCCGGCCTGGACGTAGTGGT
TGGTGGCGTCCGGACTGCTCAGGCTGAGGAAGAAGGTCAGGTAGAGCAGGCCGAGGCTGGCCGCCGCCGCGATGACCAGG
CCGATGCGCCCGCGCAGCAGGATGTTGGCAATGGCCACCGCCACCACCAGCAGGCTGCCGATGCCGCTGGGTACGCCGCC
ACCTGCGTAGAACAGGCCGCAAAGCATCAGCACGTCGGTGAGCGCGAGGATGAAGATCGGCAGCAATTGCCGCGACGGCG
GCAGGAACAGCGCGACCAGGATGTTGAAGACCAGGTAGCACCAACTGCCGACATGGAACAGTTCAGGGTGGACGAGCTTG
AGGACCTGATCTTCCAGTTCGCTGGAGATCAGCAGGACCAGTACCAGGCCGATGGTCAGGCGGTACAGGTGGTACAGACG
GAGGATGCGTTGCCCCTGCTCCTCGCTCAGCCGTAG

>ORF29729 (SEQ ID NO:334)
ACTGTTCCATGTCGGCAGTTGGTGCTACCTGGTCTTCAACATCCTGGTCGCGCTGTTCCTGCCGCCGTCGCGGCAATTGC
TGCCGATCTTCATCCTCGCGCTCACCGACGTGCTGATGCTTTGCGGCCTGTTCTACGCAGGTGGCGGCGTACCCAGCGGC
ATCGGCAGCCTGCTGGTGGTGGCGGTGGCCATTGCCAACATCCTGCTGCGCGGGCGCATCGGCCTGGTCATCGCGGCGGC
GGCCAGCCTCGGCCTGCTCTACCTGACCTTCTTCCTCAGCCTGAGCAGTCCGGACGCCACCAACCACTACGTCCAGGCCG
GCGGCCTCGGCACCCTGTGCTTCGCCGCCGCGCTGGTGATCCAGGCTCTGGTGCGGCGCCAGGAGCAGACCGAAACGCTG
GCCGAAGAACGCGCCGAGACGGTCGCCAACCTGGAGGAACTCAACGCATTGATCCTGCAGCGCATGCGCACCGGCATCCT
CGTGGTCGATAG

>ORF30221 (SEQ ID NO:336)
CCGTCAGGCCATCCTCCTCGCCAACCAGGCCGCCCTCGGCCTGCTCAGGCAGGACGACGTGCAGGGCGCCAGCCTCGGCC
GCCACAGCCCGATGCTGATGCACTGCATGAAGCAATGGCGCCTGAATCCCAGCCTCCGTCCGCCGACGCTCAAGGTGGTG
CCGGATGGCCCGACGGTGCAACCCAGCTTTATCAGCCTCAACCGCGAAGACGACCAGCACGTGCTGATCTTCCTCGAAGA
CATTTCGCAGATCGCCCAGCAGGCGCAGCAGATGAAGCTGGCCGGTCTTGGCCGCCTGACCGCCGGCATCGCCCATGA

>ORF30736c (SEQ ID NO:338)
AGCCACTCCTTCAGGTCGAGCTGCTGCGGTTCGGCCTGGCGGCGACGGGAGAGCTGCAGGACGTTCTCGATGACCAGGTT
CATCCGCTTCGACTGGTCCTGGATGATCTGCGTCAGGCGTCGGTCCGGGGCATCCAGTTCCTCTGACTCCTGCAGCAGTT
GGGCGGCGTGGCTGATCGCGCCCAGCGGGTTGCGGATCTCATGGGCGATGCCGGCGGTCAGGCGGCCAAGACCGGCCAGC
TTCATCTGCTGCGCCTGCTGGGCGATCTGCGAAATGTCTTCGAGGAAGATCAGCACGTGCTGGTCGTCTTCGCGGTTGAG
GCTGATAAAGCTGGGTTGCACCGTCGGGCCATCCGGCACCACCTTGAGCGTCGGCGGACGGAGGCTGGGATTCAGGCGCC
ATTGCTTCATGCAGTGCATCAGCATCGGGCTGTGGCGGCCGAGGCTGGCGCCCTGCACGTCGTCCTGCCTGAGCAGGCCG
AGGGCGGCCTGGTTGGCGAGGAGGATGGCCTGA

>ORF30539 (SEQ ID NO:340)
GATCCGCAACCCGCTGGGCGCGATCAGCCACGCCGCCCAACTGCTGCAGGAGTCAGAGGAACTGGATGCCCCGGACCGAC
GCCTGACGCAGATCATCCAGGACCAGTCGAAGCGGATGAACCTGGTCATCGAGAACGTCCTGCAGCTCTCCCGTCGCCGC
CAGGCCGAACCGCAGCAGCTCGACCTGAAGGAGTGGCTTCAGCGGTTCGTCGACGAATACCCCGGCAGGCTGCGCAACGA
CAGCCAACTGCACCTGCAGCTCGGTGCCGGCGACATCCAGACCCGCATGGACCCACACCAGTTGAACCAGGTGCTGAGCA
ACCTGGTGCAGAACGGTCTTCGCTACAGCGCCCAGGCGCACGGGCGCGGCCAGGTCTGGCTGAGCCTCGCGCGCGACCCG
GAGAGCGACCTGCCGGTGCTGGAAGTCATCGACGACGGTCCCGGCGTACCGGCGGACAAACTGAACAACCTGTTCGAACC
CTTCTTTACTACAGAAAGCAAAGGCACCGGCCTGGGCCTCTATCTCTCCCGCGAACTCTGCGAGAGCAACCAGGCACGGA
TCGACTACCGCAATCGCGAGGAAGGCGGCGGCTGCTTCCGCATCACCTTCGCCCACCCGCGCAAACTCAGCTGACGGAAG
CCGCACGCATGAGCCGACAAAAAGCCCTGATCGTCGACGATGAACCGGATATCCGCGAACTGCTGGAAATCACTCTCGGC
CGCATGAAGCTGGACACCCGCAGCGCCCGCAACGTCAAGGAAGCCGCGAGTTGCTGGCCCGCGAGCCGTTCGACCTGTGC
CTCACCGACATGCGCCTGCCGGACGGCAGCGGCCTCGATCTGGTCCAGTACATCCAGCAGCGCCATCCACAGACCCCGGT
GGCCATGA

<div style="text-align:center">Fig. 3-28</div>

>ORF31247c (SEQ ID NO:342)
TTTCCAGCAGTTCGCGGATATCCGGTTCATCGTCGACGATCAGGGCTTTTTGTCGGCTCATGCGTGCGGCTTCCGTCAGC
TGAGTTTGCGCGGGTGGGCGAAGGTGATGCGGAAGCAGCCGCCGCCTTCCTCGCGATTGCGGTAGTCGATCCGTGCCTGG
TTGCTCTCGCAGAGTTCGCGGGAGAGATAGAGGCCCAGGCCGGTGCCTTTGCTTTCTGTAGTAAAGAAGGGTTCGAACAG
GTTGTTCAGTTTGTCCGCCGGTACGCCGGGACCGTCGTCGATGACTTCCAGCACCGGCAGGTCGCTCTCCGGGTCGCGCG
CGAGGCTCAGCCAGACCTGGCCGCGCCCGTGCGCCTGGGCGCTGTAGCGAAGACCGTTCTGCACCAGGTTGCTCAGCACC
TGGTTCAACTGGTGTGGGTCCATGCGGGTCTGGATGTCGCCGGCACCGAGCTGCAGGTGCAGTTGGCTGTCGTTGCGCAG
CCTGCCGGGGTATTCGTCGACGAACCGCTGAAGCCACTCCTTCAGGTCGAGCTGCTGCGGTTCGGCCTGGCGGCGACGGG
AGAGCTGCAGGACGTTCTCGATGACCAGGTTCATCCGCTTCGACTGGTCCTGGATGATCTGCGTCAGGCGTCGGTCCGGG
GCATCCAGTTCCTCTGA

>ORF30963c (SEQ ID NO:344)
CTTCCAGCACCGGCAGGTCGCTCTCCGGGTCGCGCGCGAGGCTCAGCCAGACCTGGCCGCGCCCGTGCGCCTGGGCGCTG
TAGCGAAGACCGTTCTGCACCAGGTTGCTCAGCACCTGGTTCAACTGGTGTGGGTCCATGCGGGTCTGGATGTCGCCGGC
ACCGAGCTGCAGGTGCAGTTGGCTGTCGTTGCGCAGCCTGCCGGGGTATTCGTCGACGAACCGCTGAAGCCACTCCTTCA
GGTCGAGCTGCTGCGGTTCGGCCTGGCGGCGACGGGAGAGCTGCAGGACGTTCTCGATGA

>ORF31539c (SEQ ID NO:346)
GGCGGTTGCCACCAGCTCCCGCAAGCGACCGAGGTCGACCGGTTTGGTGAGGAAGTCGAAGGCACCGGCCTTGAGCGCCT
GGATCGCGGTGTCCAGGCTGCCGTACGCGGTGATCATGGCCACCGGGGTCTGTGGATGGCGCTGCTGGATGTACTGGACC
AGATCGAGGCCGCTGCCGTCCGGCAGGCGCATGTCGGTGAGGCACAGGTCGAACGGCTCGCGGGCCAGCAACTCGCGGCT
TCCTTGACGTTGCGGGCGCTGCGGGTGTCCAGCTTCATGCGGCCGAGAGTGATTTCCAGCAGTTCGCGGGATATCCGGTTC
ATCGTCGACGATCAGGGCTTTTTGTCGGCTCATGCGTGCGGCTTCCGTCAGCTGA

>ORF31222 (SEQ ID NO:348)
ACCGGATATCCGCGAACTGCTGGAAATCACTCTCGGCCGCATGAAGCTGGACACCCGCAGCGCCCGCAACGTCAAGGAAG
CCGCGAGTTGCTGGCCCGCGAGCCGTTCGACCTGTGCCTCACCGACATGCGCCTGCCGGACGGCAGCGGCCTCGATCTGG
TCCAGTACATCCAGCAGCGCCATCCACAGACCCCGGTGGCCATGATCACCGCGTACGGCAGCCTGGACACCGCGATCCAG
GCGCTCAAGGCCGGTGCCTTCGACTTCCTCACCAAACCGGTCGACCTCGGTCGCTTGCGGGAGCTGGTGGCAACCGCCCT
ACGCTTGCGCAACCCGGAAGCCGAGGAAGCGCCGGTGGACAACCGCCTGCTCGGCGAGTCGCCGCCGATGCGCGCCCTGC
GCAACCAGATCGGCAAGCTGGCGCGCAGCCAGGCGCCGGTCTACATCAGTGGCGAGTCCGGCAGCGGCAAGGAACTGGTG
GCGCGCCTGATCCACGAGCAGGGGCCACGTATCGAGCGGCCGTTCGTGCCGGTGAACTGCGGCGCGATTCCCTCCGAGCT
GATGGAAAGCGAGTTCTTCGGCCACAAGAAAGGCAGCTTCACTGGCGCTATCGAAGACAAGCAGGGCCTGTTCCAGGCCG
CCAGCGGTGGCACCCTGTTCCTCGACGAAGTCGCCGACCTGCCGATGGCCATGCAGGTCAAACTGCTCCGGGCGATCCAG
GAAAAGGCCGTGCGCGCGGTCGGCGGCCAGCAGGAGGTCGCCGTCGCACGTGCGCATCCTCTGCGCCACCCACAAGGACC
TCGCCGCCGAAGTCGGCGCCGGGCGCTTCCGCCAGGACCTCTACTACCGCCTCAACGTCATCGAGCTGCGCGTACACCGC
TGCGCGAACGCCGCGAGGACATCCCGCTGCTCGCCGAACGCATCCTCAAGCGCCTGGCCGGCGACACCGGCCTGCCGGCC
GCCAGGCTGACCGGCGACGCACAGGAGAAGCTGAAGAACTACCGCTTCCCGGGCAACGTCCGCGAGCTGGAAAACATGCT
GGAGCGCGCCTATACCCTGTGCGAAGACGACCAGATCCAGCCTCACGACCTGCGCCTGGCCGATGCGCCGGGTGCCAGCC
AGGAAGGCGCCGCGAGCCTGAGCGAAATCGACAACCTCGAGGACTACCTGGAAGACATCGAGCGCAAGCTGATCATGCAG
GCACTCGAGGAGACCCGCTGGAACCGCACCGCCGCGGCCCAGCGCCTGGGCCTGACGTTCCGCTCGATGCGCTACCGCCT
GAAAAAGCTGGGCATCGACTGA

Fig. 3-29

>ORF31266 (SEQ ID NO:350)
AGCTGGACACCCGCAGCGCCCGCAACGTCAAGGAAGCCGCGAGTTGCTGGCCCGCGAGCCGTTCGACCTGTGCCTCACCG
ACATGCGCCTGCCGGACGGCAGCGGCCTCGATCTGGTCCAGTACATCCAGCAGCGCCATCCACAGACCCCGGTGGCCATG
ATCACCGCGTACGGCAGCCTGGACACCGCGATCCAGGCGCTCAAGGCCGGTGCCTTCGACTTCCTCACCAAACCGGTCGA
CCTCGGTCGCTTGCGGGAGCTGGTGGCAACCGCCCTACGCTTGCGCAACCCGGAAGCCGAGGAAGCGCCGGTGGACAACC
GCCTGCTCGGCGAGTCGCCGCCGATGCGCGCCCTGCGCAACCAGATCGGCAAGCTGGCGCGCAGCCAGGCGCCGGTCTAC
ATCAGTGGCGAGTCCGGCAGCGGCAAGGAACTGGTGGCGCGCCTGATCCACGAGCAGGGGCCACGTATCGAGCGGCCGTT
CGTGCCGGTGAACTGCGGCGCGATTCCCTCCGAGCTGATGGAAAGCGAGTTCTTCGGCCACAAGAAAGGCAGCTTCACTG
GCGCTATCGAAGACAAGCAGGGCCTGTTCCAGGCCGCCAGCGGTGGCACCCTGTTCCTCGACGAAGTCGCCGACCTGCCG
ATGGCCATGCAGGTCAAACTGCTCCGGGCGATCCAGGAAAAGGCCGTGCGCGCGGTCGGCGGCCAGCAGGAGGTCGCCGT
CGCACGTGCGCATCCTCTGCGCCACCCACAAGGACCTCGCCGCCGAAGTCGGCGCCGGGCGCTTCCGCCAGGACCTCTAC
TACCGCCTCAACGTCATCGAGCTGCGCGTACACCGCTGCGCGAACGCCGCGAGGACATCCCGCTGCTCGCCGAACGCATC
CTCAAGCGCCTGGCCGGCGACACCGGCCTGCCGGCCGCCAGGCTGACCGGCGACGCACAGGAGAAGCTGAAGAACTACCG
CTTCCCGGGCAACGTCCGCGAGCTGGAAAAACATGCTGGAGCGCGCCTATACCCTGTGCGAAGACGACCAGATCCAGCCTC
ACGACCTGCGCCTGGCCGATGCGCCGGGTGCCAGCCAGGAAGGCGCCGCGAGCCTGAGCGAAATCGACAACCTCGAGGAC
TACCTGGAAGACATCGAGCGCAAGCTGATCATGCAGGCACTCGAGGAGACCCGCTGGAACCGCACCGCCGCGGCCCAGCG
CCTGGGCCTGACGTTCCGCTCGATGCGCTACCGCCTGAAAAAGCTGGGCATCGACTGAAAGTGAAAAGGCCTGTCCGAAG
ACAGGCCTTTTGGTTTTCGCTCCTCAGAGGCGACCAGCCGGGGCGTAGGGGGCCGGGTCGATGA

>ORF31661c (SEQ ID NO:352)
ACCGGCGCCTGGCTGCGCGCCAGCTTGCCGATCTGGTTGCGCAGGGCGCGCATCGGCGGCGACTCGCCGAGCAGGCGGTT
GTCCACCGGCGCTTCCTCGGCTTCCGGGTTGCGCAAGCGTAGGGCGGTTGCCACCAGCTCCCGCAAGCGACCGAGGTCGA
CCGGTTTGGTGAGGAAGTCGAAGGCACCGGCCTTGAGCGCCTGGATCGCGGTGTCCAGGCTGCCGTACGCGGTGATCATG
GCCACCGGGGTCTGTGGATGGCGCTGCTGGATGTACTGGACCAGATCGAGGCCGCTGCCGTCCGGCAGGCGCATGTCGGT
GAGGCACAGGTCGAACGGCTCGCGGGCCAGCAACTCGCGGCTTCCTTGA

>ORF32061c (SEQ ID NO:354)
AGGTCCTGGCGGAAGCGCCCGGCGCCGACTTCGGCGGCGAGGTCCTTGTGGGTGGCGCAGAGGATGCGCACGTGCGACGG
CGACCTCCTGCTGGCCGCCGACCGCGCGCACGGCCTTTTCCTGGATCGCCCGGAGCAGTTTGACCTGCATGGCCATCGGC
AGGTCGGCGACTTCGTCGAGGAACAGGGTGCCACCGCTGGCGGCCTGGAACAGGCCCTGCTTGTCTTCGATAGCGCCAGT
GAAGCTGCCTTTCTTGTGGCCGAAGAACTCGCTTTCCATCAGCTCGGAGGGAATCGCGCCGCAGTTCACCGGCACGAACG
GCCGCTCGATACGTGGCCCCTGCTCGTGGATCAGGCGCGCCACCAGTTCCTTGCCGCTGCCGGACTCGCCACTGATGTAG
ACCGGCGCCTGGCTGCGCGCCAGCTTGCCGATCTGGTTGCGCAGGGCGCGCATCGGCGGCGACTCGCCGAGCAGGCGGTT
GTCCACCGGCGCTTCCTCGGCTTCCGGGTTGCGCAAGCGTAG

>ORF32072c (SEQ ID NO:356)
GGCGGTAGTAGAGGTCCTGGCGGAAGCGCCCGGCGCCGACTTCGGCGGCGAGGTCCTTGTGGGTGGCGCAGAGGATGCGC
ACGTGCGACGGCGACCTCCTGCTGGCCGCCGACCGCGCGCACGGCCTTTTCCTGGATCGCCCGGAGCAGTTTGACCTGCA
TGGCCATCGGCAGGTCGGCGACTTCGTCGAGGAACAGGGTGCCACCGCTGGCGGCCTGGAACAGGCCCTGCTTGTCTTCG
ATAGCGCCAGTGAAGCTGCCTTTCTTGTGGCCGAAGAACTCGCTTTCCATCAGCTCGGAGGGAATCGCGCCGCAGTTCAC
CGGCACGAACGGCCGCTCGATACGTGGCCCCTGCTCGTGGATCAGGCGCGCCACCAGTTCCTTGCCGCTGCCGGACTCGC
CACTGATGTAG

>ORF31784 (SEQ ID NO:358)
TGGAAAGCGAGTTCTTCGGCCACAAGAAAGGCAGCTTCACTGGCGCTATCGAAGACAAGCAGGGCCTGTTCCAGGCCGCC
AGCGGTGGCACCCTGTTCCTCGACGAAGTCGCCGACCTGCCGATGGCCATGCAGGTCAAACTGCTCCGGGCGATCCAGGA
AAAGGCCGTGCGCGCGGTCGGCGGCCAGCAGGAGGTCGCCGTCGCACGTGCGCATCCTCTGCGCCACCCACAAGGACCTC
GCCGCCGAAGTCGGCGCCGGGCGCTTCCGCCAGGACCTCTACTACCGCCTCAACGTCATCGAGCTGCGCGTACACCGCTG
CGCGAACGCCGCGAGGACATCCCGCTGCTCGCCGAACGCATCCTCAAGCGCCTGGCCGGCGACACCGGCCTGCCGGCCGC
CAGGCTGA

Fig. 3-30

>ORF32568c (SEQ ID NO:360)
GGAGCGAAAACCAAAAGGCCTGTCTTCGGACAGGCCTTTTTCACTTTCAGTCGATGCCCAGCTTTTTTCAGGCGGTAGCGCA
TCGAGCGGAACGTCAGGCCCAGGCGCTGGGCCGCGGCGGTGCGGTTCCAGCGGGTCTCCTCGAGTGCCTGCATGATCAGC
TTGCGCTCGATGTCTTCCAGGTAGTCCTCGAGGTTGTCGATTTCGCTCAGGCTCGCGGCGCCTTCCTGGCTGGCACCCGG
CGCATCGGCCAGGCGCAGGTCGTGAGGCTGGATCTGGTCGTCTTCGCACAGGGTATAGGCGCGCTCCAGCATGTTTTCCA
GCTCGCGGACGTTGCCCGGGAAGCGGTAGTTCTTCAGCTTCTCCTGTGCGTCGCCGGTCAGCCTGGCGGCCGGCAGGCCG
GTGTCGCCGGCCAGGCGCTTGAGGATGCGTTCGGCGAGCAGCGGGATGTCCTCGCGGCGTTCGCGCAGCGGTGTACGCGC
AGCTCGATGACGTTGAGGCGGTAG

>ORF33157c (SEQ ID NO:362)
ACAGACGGAGGTGCGCGGCTGGTTGCGCGACGGCGATCGAGTGGTCGGCGTGGCGACCTCGCGTGGCGAGATCCGTGGCG
ACAAGGTGCTGCTGGCGGCAGGCGCCTGGAGCGGCGAGTTGTTGAAGCCGCTTGGCCTGGAACTGCCCGTGGTACCGGTG
AAAGGTCAGATGATCCTCTACAAGTGCGCGGCGGATTTCCTGCCGCGCATGGTGCTGGCCAAGGGGCGCTACGCGATTCC
GCGGCGCGACGGCCACATCCTGATCGGCAGCACCTTGGAACATTCGGGCTTCGACAAGACGCCGACCGACGAGGCGCTGG
AAAGCCTCAGGGCGTCTGCGGCAGAACTGTTGCCGGAACTGGCGGACATGCAGCCGGTGGCCCACTGGGCAGGGTTGCGC
CCGGGCTCTCCCGAAGGCATCCCCTATATCGGTCCGGTGCCTGGCTTCGACGGGCTCTGGCTGAATACCGGGCACTACCG
CAACGGGCTGGTCCTGGCACCGGCGTCGTGCCGTCTGCTGGCGGATCTCATGAGCGGGCGGGAACCGATCATCGACCCGG
CCCCCTACGCCCCGGCTGGTCGCCTCTGAGGAGCGAAAACCAAAAGGCCTGTCTTCGGACAGGCCTTTTTCACTTTCAGTC
GATGCCCAGCTTTTTTCAGGCGGTAGCGCATCGAGCGGAACGTCAGGCCCAGGCGCTGGGCCGCGGCGGTGCGGTTCCAGC
GGGTCTCCTCGAGTGCCTGCATGA

>ORF32530 (SEQ ID NO:364)
AAAGGCCTGTCCGAAGACAGGCCTTTTGGTTTTCGCTCCTCAGAGGCGACCAGCCGGGGCGTAGGGGGCCGGGTCGATGA
TCGGTTCCCGCCCGCTCATGAGATCCGCCAGCAGACGGCACGACGCCGGTGCCAGGACCAGCCCGTTGCGGTAGTGCCCG
GTATTCAGCCAGAGCCCGTCGAAGCCAGGCACCGGACCGATATAGGGGATGCCTTCGGGAGAGCCCGGGCGCAACCCTGC
CCAGTGGGCCACCGGCTGCATGTCCGCCAGTTCCGGCAACAGTTCTGCCGCAGACGCCCTGAGGCTTTCCAGCGCCTCGT
CGGTCGGCGTCTTGTCGAAGCCCGAATGTTCCAAGGTGCTGCCGATCAGGATGTGGCCGTCGCGCCGCGGAATCGCGTAG
CGCCCCTTGGCCAGCACCATGCGCGGCAGGAAATCCGCCGCGCACTTGTAGAGGATCATCTGACCTTTCACCGGTACCAC
GGGCAGTTCCAGGCCAAGCGGCTTCAACAACTCGCCGCTCCAGGCGCCTGCCGCCAGCAGCACCTTGTCGCCACGGATCT
CGCCACGCGAGGTCGCCACGCCGACCACTCGATCGCCGTCGCGCAACCAGCCGCGCACCTCCGTCTGTTCATGCAACTCG
AGATTGGCGAATTGTTGCAGGGATGCCCGCAATGA

>ORF33705c (SEQ ID NO:366)
GTGATATTTCTCTGTTCCTGGCAAATCGGTAGGAGCCCTGTGGTGAGTAGAGATGTAGTAGTGGTAGGCGCTGGCGTCAT
CGGCCTGTTGACCGCCCGGGAGCTGGCGCTCGCCGGACTGCGGGTGACCCTGGTGGAGCGGGGCGAGAGTGGGCGTGAGG
CATCCTGGGCGGGAGGCGGGATCGTCTCGCCGCTCTATCCGTGGCGCTACAGCCCGGCGGTGACCGCCCTGGCGCACTGG
TCGCAGGACTTCTACCCGGCCCTGGGGCAGCGTTTGCTCGACGAGACCGGGCTCGATCCCGAGGTCCATACCGTTGGCCT
GTACTGGCTGGACCTGGACGACCAGACCGAGGCACTGCAGTGGGCACGCAACCACACCCGGCCGTTGAAGGAAGTGCCGA
TCGAGGAGGCCTACGCGGCGGTGCCCGGGCTGGGCGCAGGCTTCCAGCGGGCGGTCTACATGTCGGGCGTGGCCAATGTG
CGCAATCCTCGCCTGGCGCGCTCATTGCGGGCATCCCTGCAACAATTCGCCAATCTCGAGTTGCATGAACAGACGGAGGT
GCGCGGCTGGTTGCGCGACGGCGATCGAGTGGTCGGCGTGGCGACCTCGCGTGGCGAGATCCGTGGCGACAAGGTGCTGC
TGGCGGCAGGCGCCTGGAGCGGCGAGTTGTTGAAGCCGCTTGGCCTGGAACTGCCCGTGGTACCGGTGAAAGGTCAGATG
ATCCTCTACAAGTGCGCGGCGGATTTCCTGCCGCGCATGGTGCTGGCCAAGGGGCGCTACGCGATTCCGCGGCGCGACGG
CCACATCCTGATCGGCAGCACCTTGGAACATTCGGGCTTCGACAAGACGCCGACCGACGAGGCGCTGGAAAGCCTCAGGG
CGTCTGCGGCAGAACTGTTGCCGGAACTGGCGGACATGCAGCCGGTGGCCCACTGGGCAGGGTTGCGCCCGGGCTCTCCC
GAAGGCATCCCCTATATCGGTCCGGTGCCTGGCTTCGACGGGCTCTGGCTGAATACCGGGCACTACCGCAACGGGCTGGT
CCTGGCACCGGCGTCGTGCCGTCTGCTGGCGGATCTCATGAGCGGGCGGGAACCGATCATCGACCCGGCCCCCTACGCCC
CGGCTGGTCGCCTCTGA

Fig. 3-31

>ORF32832 (SEQ ID NO:368)
GGCTTTCCAGCGCCTCGTCGGTCGGCGTCTTGTCGAAGCCCGAATGTTCCAAGGTGCTGCCGATCAGGATGTGGCCGTCG
CGCCGCGGAATCGCGTAGCGCCCCTTGGCCAGCACCATGCGCGGCAGGAAATCCGCCGCGCACTTGTAGAGGATCATCTG
ACCTTTCACCGGTACCACGGGCAGTTCCAGGCCAAGCGGCTTCAACAACTCGCCGCTCCAGGCGCCTGCCGCCAGCAGCA
CCTTGTCGCCACGGATCTCGCCACGCGAGGTCGCCACGCCGACCACTCGATCGCCGTCGCGCAACCAGCCGCGCACCTCC
GTCTGTTCATGCAACTCGAGATTGGCGAATTGTTGCAGGGATGCCCGCAATGAGCGCGCCAGGCGAGGATTGCGCACATT
GGCCACGCCCGACATGTAGACCGCCCGCTGGAAGCCTGCGCCCAGCCCGGGCACCGCCGCGTAGGCCTCCTCGATCGGCA
CTTCCTTCAACGGCCGGGTGTGGTTGCGTGCCCACTGCAGTGCCTCGGTCTGGTCGTCCAGGTCCAGCCAGTACAGGCCA
ACGGTATGGACCTCGGGATCGAGCCCGGTCTCGTCGAGCAAACGCTGCCCCAGGGCCGGGTAGAAGTCCTGCGACCAGTG
CGCCAGGGCGGTCACCGCCGGGCTGTAGCGCCACGGATAGAGCGGCGAGACGATCCCGCCTCCCGCCCAGGATGCCTCAC
GCCCACTCTCGCCCCGCTCCACCAGGGTCACCCGCAGTCCGGCGAGCGCCAGCTCCCGGGCGGTCAACAGGCCGATGA

>ORF33547c (SEQ ID NO:370)
GGCATCCTGGGCGGGAGGCGGGATCGTCTCGCCGCTCTATCCGTGGCGCTACAGCCCGGCGGTGACCGCCCTGGCGCACT
GGTCGCAGGACTTCTACCCGGCCCTGGGGCAGCGTTTGCTCGACGAGACCGGGCTCGATCCCGAGGTCCATACCGTTGGC
CTGTACTGGCTGGACCTGGACGACCAGACCGAGGCACTGCAGTGGGCACGCAACCACACCCGGCCGTTGAAGGAAGTGCC
GATCGAGGAGGCCTACGCGGCGGTGCCCGGGCTGGGCGCAGGCTTCCAGCGGGCGGTCTACATGTCGGGCGTGGCCAATG
TGCGCAATCCTCGCCTGGCGCGCTCATTGCGGGCATCCCTGCAACAATTCGCCAATCTCGAGTTGCATGA

>ORF33205 (SEQ ID NO:372)
GCGCGCCAGGCGAGGATTGCGCACATTGGCCACGCCCGACATGTAGACCGCCCGCTGGAAGCCTGCGCCCAGCCCGGGCA
CCGCCGCGTAGGCCTCCTCGATCGGCACTTCCTTCAACGGCCGGGTGTGGTTGCGTGCCCACTGCAGTGCCTCGGTCTGG
TCGTCCAGGTCCAGCCAGTACAGGCCAACGGTATGGACCTCGGGATCGAGCCCGGTCTCGTCGAGCAAACGCTGCCCCAG
GGCCGGGTAGAAGTCCTGCGACCAGTGCGCCAGGGCGGTCACCGCCGGGCTGTAGCGCCACGGATAGAGCGGCGAGACGA
TCCCGCCTCCCGCCCAGGATGCCTCACGCCCACTCTCGCCCCGCTCCACCAGGGTCACCCGCAGTCCGGCGAGCGCCAGC
TCCCGGGCGGTCAACAGGCCGATGACGCCAGCGCCTACCACTACTACATCTCTACTCACCACAGGGCTCCTACCGATTTG
CCAGGAACAGAGAAATATCACTCAAAGGGATCAGATGCTGACGAATTGCCTGCTTCAACGAACTCAGTCGAATCTAGTCC
CGGTGAAAAGCCCATCATACCCGCAGAGGTATTCATCCCATGA

>ORF33512 (SEQ ID NO:374)
AGCGGCGAGACGATCCCGCCTCCCGCCCAGGATGCCTCACGCCCACTCTCGCCCCGCTCCACCAGGGTCACCCGCAGTCC
GGCGAGCGCCAGCTCCCGGGCGGTCAACAGGCCGATGACGCCAGCGCCTACCACTACTACATCTCTACTCACCACAGGGC
TCCTACCGATTTGCCAGGAACAGAGAAATATCACTCAAAGGGATCAGATGCTGACGAATTGCCTGCTTCAACGAACTCAG
TCGAATCTAGTCCCGGTGAAAAGCCCATCATACCCGCAGAGGTATTCATCCCATGAAATCGAGTGGTTTGAATTTGGTGG
AACTATCGATAGTCCTATCGATCCTTGCGATAGGCGTGACAATTGCGCTGCCCACCCTCCCCGACAGAATGAAGCGGGAC
ATTAG

>ORF33771 (SEQ ID NO:376)
AAAGCCCATCATACCCGCAGAGGTATTCATCCCATGAAATCGAGTGGTTTGAATTTGGTGGAACTATCGATAGTCCTATC
GATCCTTGCGATAGGCGTGACAATTGCGCTGCCCACCCTCCCCGACAGAATGAAGCGGGACATTAGCCGTGATATTGGTG
ACAGCCTGACTAGTCATGTGATGGCTGCGCGGGCTAGCAGCATACAGAACGGCGTGATCATCGAGGTGTGCGGTAGCGGT
GACGGCAGTACCTGCAGCGAGGAATGGCATCTCGGCTGGTTCAGCCGTAACGACAGGAGCCAACAGATACTGGCCCGGCA
TGAAAATACGAGTCGCACCGATATTCATTGGCGGGGCTTCGACAAGCGACTGCGCTACCTGCCTAATGGCACCAGCCCTA
CAGGTAACGGGCGTTTCTTCGAATGTAAGGACGATCGCATCGAGTGGCAATTGGTGCTCAATCGGCAAGGCCGCCTCAGG
GTGGCGGGAAAGAGCGAAAATAAAAAGCTCTCTTACCTGTGCTCCAGGCGGTGA

>ORF34385c (SEQ ID NO:378)
TGGAGAGCGCATTGTCCCTGTAGCAGAGACAGCCGGAGCGGAGAGTGGGATGACTGGCAAACGGTATGTGAAACAGTTCT
CTCACCGCCTGGAGCACAGGTAAGAGAGCTTTTTATTTTCGCTCTTTCCCGCCACCCTGAGGCGGCCTTGCCGATTGAGC
ACCAATTGCCACTCGATGCGATCGTCCTTACATTCGAAGAAACGCCCGTTACCTGTAGGGCTGGTGCCATTAGGCAGGTA
GCGCAGTCGCTTGTCGAAGCCCCGCCAATGAATATCGGTGCGACTCGTATTTTCATGCCGGGCCAGTATCTGTTGGCTCC
TGTCGTTACGGCTGAACCAGCCGAGATGCCATTCCTCGCTGCAGGTACTGCCGTCACCGCTACCGCACACCTCGATGATC
ACGCCGTTCTGTATGCTGCTAGCCCGCGCAGCCATCACATGACTAGTCAGGCTGTCACCAATATCACGGCTAATGTCCCG
CTTCATTCTGTCGGGGAGGGTGGGCAGCGCAATTGTCACGCCTATCGCAAGGATCGATAG

Fig. 3-32

>ORF33988 (SEQ ID NO:380)
TCATCGAGGTGTGCGGTAGCGGTGACGGCAGTACCTGCAGCGAGGAATGGCATCTCGGCTGGTTCAGCCGTAACGACAGG
AGCCAACAGATACTGGCCCGGCATGAAAATACGAGTCGCACCGATATTCATTGGCGGGGCTTCGACAAGCGACTGCGCTA
CCTGCCTAATGGCACCAGCCCTACAGGTAACGGGCGTTTCTTCGAATGTAAGGACGATCGCATCGAGTGGCAATTGGTGC
TCAATCGGCAAGGCCGCCTCAGGGTGGCGGGAAAGAGCGAAAATAAAAAGCTCTCTTACCTGTGCTCCAGGCGGTGAGAG
AACTGTTTCACATACCGTTTGCCAGTCATCCCACTCTCCGCTCCGGCTGTCTCTGCTACAGGGACAATGCGCTCTCCACT
AG

>ORF34274 (SEQ ID NO:382)
AAAGCTCTCTTACCTGTGCTCCAGGCGGTGAGAGAACTGTTTCACATACCGTTTGCCAGTCATCCCACTCTCCGCTCCGG
CTGTCTCTGCTACAGGGACAATGCGCTCTCCACTAGGCAAGATTATCTGGCCCTTTTCCTTGTGGAGTACTGCATGCGCT
CTATTTGTCGCAGCGCCGGCTTTTCCCTGATCGAGTTGATGATGGTGTTGGTTCTGGTCGCCATATTCGCCAGCATTGCC
GTACCCAGTTTCAACGCCTTGATCGAGCGCAACCGAATCCAGACTGCCAGCGAGGAACTCTACAGCCTGCTTCAGTACGC
TCGCAGCGAAGCTGTAAACCGTCATGCCAATGTGAGCATCAGGGCGACGCAGAACAATGACTGGGCAAAAGGCCTGGAAA
TCATCAGCGGCGCGACCACCGTGCAAAAGCACCAAGGTTTCCAGCAGGTCTCGCTATCCGCCAGCAGTGCGACTGCGGAG
CTGACCTTCAACGCTACCGGCACACTTAGCAACCAGGCTGCAAACATTGACATAAAGGTCTGCTTCGCCGGTGACAAAAG
TACAGGACGTCTGCTTACCGTTCAGCCCAGTGGACGCGTGATCCTGTACCCATCTTCAAAGCAACCGGACAGCTGTAACT
GA

>ORF34726c (SEQ ID NO:384)
CGAGACCTGCTGGAAACCTTGGTGCTTTTGCACGGTGGTCGCGCCGCTGATGATTTCCAGGCCTTTTGCCCAGTCATTGT
TCTGCGTCGCCCTGATGCTCACATTGGCATGACGGTTTACAGCTTCGCTGCGAGCGTACTGAAGCAGGCTGTAGAGTTCC
TCGCTGGCAGTCTGGATTCGGTTGCGCTCGATCAAGGCGTTGAAACTGGGTACGGCAATGCTGGCGAATATGGCGACCAG
AACCAACACCATCATCAACTCGATCAGGGAAAAGCCGGCGCTGCGACAAATAGAGCGCATGCAGTACTCCACAAGGAAAA
GGGCCAGATAATCTTGCCTAGTGGAGAGCGCATTGTCCCTGTAGCAGAGACAGCCGGAGCGGAGAGTGGGATGACTGGCA
AACGGTATGTGAAACAGTTCTCTCACCGCCTGGAGCACAGGTAA

>ORF34916 (SEQ ID NO:386)
GGAAAGCCCATGTCTCGAGAAACGGGTTTCAGCATGATCGAAGTACTGGTTGCTCTGGTGCTGATCAGCATTGGCGTACT
GGGCATGGTTGCCATGCAAGGGCGCACGATCCAGTACACGCAGGAGTCGGTACAACGCAATGCCGCAGCAATGCTTGCTA
GCGACCTGATGGAAATAATGCGTGCGGACCCAGATGCCGTACTCAATCTACGCGCCCAACTACGCGAAGACTCGGTCTAC
TACAAGGCCAAGGGCAGCGACTTTCCCGCAGCCCCAGCGCGCTGCGCGCCATTGCCAGCAGATGCTAAGGAACGTCTCGG
CTGCTGGGCCCAACAGGCCTCGAAAGACTTGCCGGGAGCCTCCGCACTCTTGAATAGCCAATTCTACATTTGTCGCAGCC
CAACCCCGGGTACCTGCGACAACACCAAAGGCTCGGCCATCGAAATCCAGGTTGCCTGGCGAGCCATGGATGGAGCGTGT
TTCAACGCCTCTGACTCCACCTTGTGCACCTACAGCGTCCGCTCCGAATTGTGA

>ORF35464c (SEQ ID NO:388)
AGAGCATGCTTGTTCTCACAATTCGGAGCGGACGCTGTAGGTGCACAAGGTGGAGTCAGAGGCGTTGAAACACGCTCCAT
CCATGGCTCGCCAGGCAACCTGGATTTCGATGGCCGAGCCTTTGGTGTTGTCGCAGGTACCCGGGGGTTGGGCTGCGACAA
ATGTAGAATTGGCTATTCAAGAGTGCGGAGGCTCCCGGCAAGTCTTTCGAGGCCTGTTGGGCCCAGCAGCCGAGACGTTC
CTTAGCATCTGCTGGCAATGGCGCGCAGCGCGCTGGGGCTGCGGGAAAGTCGCTGCCCTTGGCCTTGTAGTAGACCGAGT
CTTCGCGTAG

>ORF35289 (SEQ ID NO:390)
ATAGCCAATTCTACATTTGTCGCAGCCCAACCCCGGGTACCTGCGACAACACCAAAGGCTCGGCCATCGAAATCCAGGTT
GCCTGGCGAGCCATGGATGGAGCGTGTTTCAACGCCTCTGACTCCACCTTGTGCACCTACAGCGTCCGCTCCGAATTGTG
AGAACAAGCATGCTCTTCAGCAAAATGCAGAAAGGCCTATCGATGGTAGAACTGCTCGTGGCACTCGCTATAAGCAGCTT
CCTGATCCTGGGGATCAGCCAGATCTACATCGACAACAAACGCAACTATCTTTTCCAGCAAGGCCAGGCCGGCAACCAGG
AAAATAG

Fig. 3-33

>ORF35410 (SEQ ID NO:392)
CTCCACCTTGTGCACCTACAGCGTCCGCTCCGAATTGTGAGAACAAGCATGCTCTTCAGCAAAATGCAGAAAGGCCTATC
GATGGTAGAACTGCTCGTGGCACTCGCTATAAGCAGCTTCCTGATCCTGGGGATCAGCCAGATCTACATCGACAACAAAC
GCAACTATCTTTTCCAGCAAGGCCAGGCCGGCAACCAGGAAAATAGCCGCTTCGTTCTTATGCTGCTGCAGCAACAACTG
GATAAGACAGCCTATCGTCGCCTTCACGACGACAACATGGAGAATGCTTTCAAATCCGCGACATTCAATGGCTGTCGTGC
ATTTGTGGCTGGCGAGACTATCGCTGCGGCAACTGCCCTCAAGGCGGGTGAGTACGGTGTCTGCTTGCGCTATCAACCCG
CCTACAAAGGGGAGCATGATTGCCTCGGTAATGAAATTACCGGAGTTCCGGAAAAGCCCTTCACAAATACTCCCCCTGTC
GTCGTTCGCCTGGTCTACCTACCGAGCGCCGGTACCCTGAGTTGCAGTCGTCCCGATATCGCCCAGTCGAAATCGGGAGA
ATTGGTCAGTGGTCTCACAGACTTCCGCTTGGAAGCGGGGGTCGGGCCAGCAGATCGTAGCGAACGCAAAGTATCCAGCT
TCGTCGCACTACAGGATGTCGCCGGTCGTCCTATCCGAGCATTGCGCTTCTCAATCCTGGCAGGCAGCGACAATACAAGC
CTGCGCACAGGAGATGATAGCCAGGCACGCGATCGCTGGATCGTCCTTTATCCCGAGAGCAAAAGCGCCATCGAGGCCGC
AGACAAAGGCCAGATTTACCAAATAGCGCGTGGTAACCAAACCATCAGGAATCTCATGCCATGA

>ORF35907c (SEQ ID NO:394)
GTAGACCAGGCGAACGACGACAGGGGGAGTATTTGTGAAGGGCTTTTCCGGAACTCCGGTAATTTCATTACCGAGGCAAT
CATGCTCCCCTTTGTAGGCGGGTTGATAGCGCAAGCAGACACCGTACTCACCCGCCTTGAGGGCAGTTGCCGCAGCGATA
GTCTCGCCAGCCACAAATGCACGACAGCCATTGAATGTCGCGGATTTGAAAGCATTCTCCATGTTGTCGTCGTGAAGGCG
ACGATAGGCTGTCTTATCCAGTTGTTGCTGCAGCAGCATAAGAACGAAGCGGCTATTTTCCTGGTTGCCGGCCTGGCCTT
GCTGGAAAAGATAGTTGCGTTTGTTGTCGATGTAGATCTGGCTGATCCCCAGGATCAGGAAGCTGCTTATAGCGAGTGCC
ACGAGCAGTTCTACCATCGATAG

>ORF35534 (SEQ ID NO:396)
TCCTGGGGATCAGCCAGATCTACATCGACAACAAACGCAACTATCTTTTCCAGCAAGGCCAGGCCGGCAACCAGGAAAAT
AGCCGCTTCGTTCTTATGCTGCTGCAGCAACAACTGGATAAGACAGCCTATCGTCGCCTTCACGACGACAACATGGAGAA
TGCTTTCAAATCCGCGACATTCAATGGCTGTCGTGCATTTGTGGCTGGCGAGACTATCGCTGCGGCAACTGCCCTCAAGG
CGGGTGAGTACGGTGTCTGCTTGCGCTATCAACCCGCCTACAAAGGGGAGCATGATTGCCTCGGTAATGAAATTACCGGA
GTTCCGGAAAAGCCCTTCACAAATACTCCCCCTGTCGTCGTTCGCCTGGTCTACCTACCGAGCGCCGGTACCCTGA

>ORF35930 (SEQ ID NO:398)
GTTGCAGTCGTCCCGATATCGCCCAGTCGAAATCGGGAGAATTGGTCAGTGGTCTCACAGACTTCCGCTTGGAAGCGGGG
GTCGGGCCAGCAGATCGTAGCGAACGCAAAGTATCCAGCTTCGTCGCACTACAGGATGTCGCCGGTCGTCCTATCCGAGC
ATTGCGCTTCTCAATCCTGGCAGGCAGCGACAATACAAGCCTGCGCACAGGAGATGATAGCCAGGCACGCGATCGCTGGA
TCGTCCTTTATCCCGAGAGCAAAAGCGCCATCGAGGCCGCAGACAAAGGCCAGATTTACCAAATAG

>ORF36246 (SEQ ID NO:400)
CCAAACCATCAGGAATCTCATGCCATGACCCTGCGCCATACCTCTCGACAGCAGGGATCCACGTTGTTGATCTCGCTGGT
TATCTTGTTGATGATCACGCTCCTCGCCGTTTCCAACATGCGCGAGGTGTCACTGGAAAGCCGTATCACCGGCAATCTCA
TCGAACAGAAGCGCCTGCGCAATGCGGGCGAAGCTGGGCTACGCGAAGGTGAACGACGCTTTTTCAATACCATCAAGCCC
CCAGAGGTCGGCAGCGGATGCGCCGATAGCAATGTCAAACGGCCTTGCATACTGAACCTGAGTGCCCTCTCCGTACCCCG
AGATGACGTGCACAACAATCCGGTGGCAGCCCTGAACGGCAAGACAGATAACGCCAATTCACGTGTCTGGATGCCCTACC
GAGGCAGCGATCTGAATAACCCTACGCAGATCGACAAAGACCGCGCAGTCACCTGGCAGACCATCACGGTGCCCGCTGGC
GAACAGAACAACGAAGCGGAAAATCCCGAGTACGGCAACATGATGCGCGGGGTCGGCACGTTCTACTACGAAACCAACAG
CCGCGCCCTCAACAAGGCGGGCGGAGAGACTGTTCTACAGGCCGTTCATGCACGCCTGTATACCAACTGA

>ORF26640c (SEQ ID NO:402)
GGCATCCAGACACGTGAATTGGCGTTATCTGTCTTGCCGTTCAGGGCTGCCACCGGATTGTTGTGCACGTCATCTCGGGG
TACGGAGAGGGCACTCAGGTTCAGTATGCAAGGCCGTTTGACATTGCTATCGGCGCATCCGCTGCCGACCTCTGGGGGCT
TGATGGTATTGAAAAAGCGTCGTTCACCTTCGCGTAGCCCAGCTTCGCCCGCATTGCGCAGGCGCTTCTGTTCGATGAGA
TTGCCGGTGATACGGCTTTCCAGTGACACCTCGCGCATGTTGGAAACGGCGAGGAGCGTGATCATCAACAAGATAACCAG
CGAGATCAACAACGTGGATCCCTGCTGTCGAGAGGTATGGCGCAGGGTCATGGCATGA

Fig. 3-34

>ORF36769 (SEQ ID NO:404)
TGCGCGGGGTCGGCACGTTCTACTACGAAACCAACAGCCGCGCCCTCAACAAGGCGGGCGGAGAGACTGTTCTACAGGCC
GTTCATGCACGCCTGTATACCAACTGACTGGAGCCAGCGCATGATCCACCAGATTACCCGCGCAGGAAAAAGCCTGCTGG
CTGCAGGTTGCACCCTGAGCATCCTGTTCGCCTCTGACAGTTATGCCGCCACGGCCCTGAATGTCAGCCAGCAACCCCTG
TTCCTAACCCAGGGCGTTGCTCCCAACCTGCTGTTCACTCTAGATGACTCAGGCAGTATGGCCTGGGCTTACGTGCCCGA
CGGTATTAGCGGGAATAGCGGCAGAGCGGGACGTTCCAGCGATTACAACGCACTGTACTACAACCCCGATTATGCTTACC
AAGTGCCCAAGAAATTGACACTGTCAGGCGATCAGATCATCGTTTCCGACTATCCAGTGCCACGCTTCACAGCAGCCTGG
CAGGATGGCTACGCCCAAGGCTCCACCACCAACCTGAGCAATAACTATCGCCCTCAATGGGGAACCGGCTGGCTTGGTTG
CATCGATAGCAGCTGCAATACCGGGAGAGCTTATTACTATACTTATAAGGTAAGCGCTAGCTGCCCTGCACAGCCGGTGA
GCAGCTCCAACTCCTGTTATACCTACAATGCTCTTCCTACCAGTCAGGAAAGCAACTTTGCGATATGGTACTCCTACTAT
CGCAACCGCATCCTGGCCACAAAGACCGCTGCCAACCTGGCCTTTTACAGCCTGCCGGAAAACGTGCGTCTCACTTGGGG
GGCCCTGAACACCTGTAGCATCGGCGCCAACAGCAGAAGCTGCCAAAACAATGCCCTGCTCCAATTCAACAAGCAGCACA
AAATCAATTTCTTCAATTGGCTGGCGAACAGCCCGGCCAGCGGCGGTACTCCTCTGCATGCGGCTCTTGACCGAGCCGGA
CGCTTCTTGCAAACCAACGGCACAGCTTATACCACCGAAGACGGAAAGACATATTCCTGCCGGGCCAGCTATCACATCAT
GATGACCGACGGTATCTGGAACGGTCGGAACGTCACCCCCGGCAATCTCGACAACCAGAACCAGACCTTTCCTGATAGCA
CCCTCTATAGGCCACAGCCCCCTTATGCCGACAGCAATGCCAGCTCATTGGCTGACCTGGCTTTCAAATACTGGACCACA
GACTTACGTCCCAGCATCGACAATGACCTGAAGCCTTTCATGGCCTACAAGAGTGGGGACGATTCCAAGGATTACTGGGA
CCCTCGCAACAACCCAGCCACTTGGCAACACATGGTCAACTTTACCGTTGGCCTAGGTCTTTCCTATTCGCTCACATTGA
ACTCTGCACCAACTTGGACAGGCAGCACCTTTGGCAACTACGAGGAGTTGATGGCTGGAAGCAAGGCTTGGCCCAGCGTC
GATAACGACGCCGCACCCGGTAACGTCTACGACCTCTGGCATGCAGCTATCAACTCTCGTGGAGACTTCTTTAGCGCGGA
ATCACCGGACTCTCTGGTTCAGGCTTTCAATAAGATCCTGACACGGATTTCCGAGCGCAACACCTCCTCCTCCAAACCAG
CAATGACTTCCGCGCTGCAGGATGACGGAACCGGCGACAAGCTGATCCGCTACAGCTACCAGTCCAGCTTTGCCAGTGAC
AAGAACTGGGCGGGCGACCTTATACGTTACAAGGTGGAGTCGACTTCCACCGGTTCGACCAAAAACCCAGGAATGGAGCGC
CGGCGCACTGCTGGACAACCGAGCTCCCGCTACCCGTAATATTTACATCGCCAGCAATAGCGGAACCAACCGCCTTAAGC
CTTTCACATGGAGCAATATTGAGGGAAGTCAGTTAGCCACTTGGCTGAACCGCAACCCGGACAAGGACAATCAGGCCGAC
ACCAAAGGAGCACAGCGGGTCGACTTCATCCGTGGCCAGCAGAATATGGATGGATTCCGGCAACGACAGGCGGTGTTAGG
GGACATCGTGCACTCGTCTCCAGCCGTGGTCGGACCGGCCCAATACCTCACTTATCTGGCCAACCCCATCGAACCCAGCG
GCGACTACGGCACATTCAAGACAGAGGCAGACCAGCGCAGCCCTAGAGTTTATGTTGGATCCAACGATGGCATGTTGCAT
GGTTTCAACATCAAAACCGGCGTGGAAGAGTTCGCTTTCATCCCTACAGCAGTATTCGAAAAGCTTAACAAGCTTACCGG
CATCAGCTACCAGGGCGGTGCCCACCAATATTTCGTCGACGCTACACCGGTCGTCAGCGATGCCTTTTTCGATGGAGCTT
GGCACACTGTTCTGATCGGAACGCTTGGTGCTGGAGGTCGCGGCCTGTTCGCACTCGATGTAACCAAGCCGGACGATGTC
AAGCTGCTTTGGGAATACGATAGCAGTACCGACTCGGACCTTGGTTACACCTTCTCCAAACCTACCGTAGCCAGACTGCA
CAGCGGACAATGGGCAGTAGTTACCGGCAACGGCTATGGAAGCGATAATGACAAGGCAGCTTTACTGCTGATTGATTTGA
AAAAGGGAACGCTGATCAAGAAGCTGGAAGTCCAAAGCGAGCGCGGAATAGCCAATGGCCTATCGACGCCTCGCCTGGCT
GATAACAACAGCGATGGCATTGCTGACTACGCCTATGCTGGCGATCTGCAGGGAAATATCTGGCGCTTCGATTTGATCGG
CAATACCCGCAACGACGACCCAGACACAAATACCTCTATCAATCCCTTCAAGCCCGGAGATGTAGATCCTTCTGCTTTCA
GAGTATCGTTCAGCGGCGCCCCGCTTTTCCGTGCTCGCGCCGACAACAATACTCGTCAGCCCATCACGGCTCCGCCTACC
TTGGTACGCCATCCTAGCCGTAAGGGCTACATCGTCATCGTAGGTACAGGAAAATACTTCGAGGACGATGACGCTCAGGC
CGATACCAGCCGAGCCATGACGCTCTATGGTATCTGGGATCGCCAGACCAAGGGCGAAAGCGCAAACAGTACCCCAACCA
TCGACCGCAACGCCCTCACAGCCCAAACCATGACAACAGAGGCGAACTCCACATTCGGTAGCGTGAACAGGAATATTCGG
CTTATTAGCCAAAACCCGGTGAAGTGGTACAAAGACGGAGCAACCGGTACCGCGAACTCGGATGTGGCTAGCTATGGCTG
GCGACTGAATCTGGAGGTCAATAGCAGCAAGAAAGGCGAAATGATGATCGAAGATATGTTCGCTGCCGGCCAAGTGCTTC
TATTGCAGACCTTGACACCGAACGACGACCCTTGTGACAGCGGCTCTACCAGCTGGACCTACGGCCTCAATCCATATACT
GGCGGACGTACCAGTTTCACCGTCTTCGATCTCAAACGTGCGGGTATAGTGGACTCTGGCTCGGATTACAACGGCTCGGT
CGTATCCGCCTTCCAACAGGATGGACTAGGTGGCTTGGCCATTACCCAGAACGAACAGCGTCAATCCGAGGCTTGCACTG
GTGATGAGTGCATCATCTTCAACCCCAGCGACAAGAGTAACGGACGACAAACCTGGCGGGTCGTCGAGGAGAAATGA

Fig. 3-35

>ORF37932c (SEQ ID NO:406)
GCTGGCATTGCTGTCGGCATAAGGGGGCTGTGGCCTATAGAGGGTGCTATCAGGAAAGGTCTGGTTCTGGTTGTCGAGAT
TGCCGGGGGTGACGTTCCGACCGTTCCAGATACCGTCGGTCATCATGATGTGATAGCTGGCCCGGCAGGAATATGTCTTT
CCGTCTTCGGTGGTATAAGCTGTGCCGTTGGTTTGCAAGAAGCGTCCGGCTCGGTCAAGAGCCGCATGCAGAGGAGTACC
GCCGCTGGCCGGGCTGTTCGCCAGCCAATTGAAGAAATTGATTTTGTGCTGCTTGTTGAATTGGAGCAGGGCATTGTTTT
GGCAGCTTCTGCTGTTGGCGCCGATGCTACAGGTGTTCAGGGCCCCCCAAGTGAGACGCACGTTTTCCGGCAGGCTGTAA
AAGGCCAGGTTGGCAGCGGTCTTTGTGGCCAGGATGCGGTTGCGATAGTAGGAGTACCATATCGCAAAGTTGCTTTCCTG
ACTGGTAGGAAGAGCATTGTAGGTATAACAGGAGTTGGAGCTGCTCACCGGCTGTGCAGGGCAGCTAGCGCTTACCTTAT
AAGTATAGTAATAAGCTCTCCCGGTATTGCAGCTGCTATCGATGCAACCAAGCCAGCCGGTTCCCCATTGAGGGCGATAG
TTATTGCTCAGGTTGGTGGTGGAGCCTTGGGCGTAGCCATCCTGCCAGGCTGCTGTGAAGCGTGGCACTGGATAGTCGGA
AACGATGATCTGATCGCCTGA

>ORF38640c (SEQ ID NO:408)
CTGACTTCCCTCAATATTGCTCCATGTGAAAGGCTTAAGGCGGTTGGTTCCGCTATTGCTGGCGATGTAAATATTACGGG
TAGCGGGAGCTCGGTTGTCCAGCAGTGCGCCGGCGCTCCATTCCTGGGTTTTGGTCGAACCGGTGGAAGTCGACTCCACC
TTGTAACGTATAAGGTCGCCCGCCCAGTTCTTGTCACTGGCAAAGCTGGACTGGTAGCTGTAGCGGATCAGCTTGTCGCC
GGTTCCGTCATCCTGCAGCGCGGAAGTCATTGCTGGTTTGGAGGAGGAGGTGTTGCGCTCGGAAATCCGTGTCAGGATCT
TATTGAAAGCCTGAACCAGAGAGTCCGGTGA

>ORF39309c (SEQ ID NO:410)
AGCTGCCTTGTCATTATCGCTTCCATAGCCGTTGCCGGTAACTACTGCCCATTGTCCGCTGTGCAGTCTGGCTACGGTAG
GTTTGGAGAAGGTGTAACCAAGGTCCGAGTCGGTACTGCTATCGTATTCCCAAAGCAGCTTGACATCGTCCGGCTTGGTT
ACATCGAGTGCGAACAGGCCGCGACCTCCAGCACCAAGCGTTCCGATCAGAACAGTGTGCCAAGCTCCATCGAAAAAGGC
ATCGCTGACGACCGGTGTAGCGTCGACGAAATATTGGTGGGCACCGCCCTGGTAGCTGATGCCGGTAAGCTTGTTAAGCT
TTTCGAATACTGCTGTAGGGATGAAAGCGAACTCTTCCACGCCGGTTTTGATGTTGAAACCATGCAACATGCCATCGTTG
GATCCAACATAAACTCTAGGGCTGCGCTGGTCTGCCTCTGTCTTGAATGTGCCGTAGTCGCCGCTGGGTTCGATGGGGTT
GGCCAGATAAGTGAGGTATTGGGCCGGTCCGACCACGGCTGGAGACGAGTGCACGATGTCCCCTAA

>ORF38768 (SEQ ID NO:412)
GGGACATCGTGCACTCGTCTCCAGCCGTGGTCGGACCGGCCCAATACCTCACTTATCTGGCCAACCCCATCGAACCCAGC
GGCGACTACGGCACATTCAAGACAGAGGCAGACCAGCGCAGCCCTAGAGTTTATGTTGGATCCAACGATGGCATGTTGCA
TGGTTTCAACATCAAAACCGGCGTGGAAGAGTTCGCTTTCATCCCTACAGCCAGTATTCGAAAAGCTTAACAAGCTTACCG
GCATCAGCTACCAGGGCGGTGCCCACCAATATTTCGTCGACGCTACACCGGTCGTCAGCGATGCCTTTTTCGATGGAGCT
TGGCACACTGTTCTGA

>ORF40047c (SEQ ID NO:414)
AAGCACTTGGCCGGCAGCGAACATATCTTCGATCATCATTTCGCCTTTCTTGCTGCTATTGACCTCCAGATTCAGTCGCC
AGCCATAGCTAGCCACATCCGAGTTCGCGGTACCGGTTGCTCCGTCTTTGTACCACTTCACCGGGTTTTGGCTAATAAGC
CGAATATTCCTGTTCACGCTACCGAATGTGGAGTTCGCCTCTGTTGTCATGGTTTGGGCTGTGAGGGCGTTGCGGTCGAT
GGTTGGGGTACTGTTTGCGCTTTCGCCCTTGGTCTGGCGATCCCAGATACCATAGAGCGTCATGGCTCGGCTGGTATCGG
CCTGAGCGTCATCGTCCTCGAAGTATTTTCCTGTACCTACGATGACGATGTAGCCCTTACGGCTAGGATGGCGTACCAAG
GTAGGCGGAGCCGTGATGGGCTGACGAGTATTGTTGTCGGCGCGAGCACGGAAAAGCGGGGCGCCGCTGAACGATACTCT
GAAAGCAGAAGGATCTACATCTCCGGGCTTGAAGGGATTGATAGAGGTATTTGTGTCTGGGTCGTCGTTGCGGGTATTGC
CGATCAAATCGAAGCGCCAGATATTTCCCTGCAGATCGCCAGCATAGGCGTAGTCAGCAATGCCATCGCTGTTGTTATCA
GCCAGGCGAGGCGTCGATAG

>ORF40560c (SEQ ID NO:416)
CCGGCGAGTCCTGTTGTTGGACACGGTTGGGCAAGCGATATGTCTGCCCATCGACTACTACCAGACCGGCGGCAGGATGA
ACATCCTCGACCACGCCCACATTCTCGAACGTATTCGTGGCACTCAAGGCAAAGGTTGGGCAAGCCAGAGCTAGAGCTGC
AAGAGCTGTGGCGAGAAGACGTAAGGGGGTTCATGTTCATTTCTCCTCGACGACCCGCCAGGTTTGTCGTCCGTTACTCTT
GTCGCTGGGGTTGAAGATGATGCACTCATCACCAGTGCAAGCCTCGGATTGACGCTGTTCGTTCTGGGTAATGGCCAAGC
CACCTAG

Fig. 3-36

>ORF40238 (SEQ ID NO:418)
GTGGCTTGGCCATTACCCAGAACGAACAGCGTCAATCCGAGGCTTGCACTGGTGATGAGTGCATCATCTTCAACCCCAGC
GACAAGAGTAACGGACGACAAACCTGGCGGGTCGTCGAGGAGAAATGAACATGAACCCCTTACGTCTTCTCGCCACAGCT
CTTGCAGCTCTAGCTCTGGCTTGCCCAACCTTTGCCTTGAGTGCCACGAATACGTTCGAGAATGTGGGCGTGGTCGAGGA
TGTTCATCCTGCCGCCGGTCTGGTAGTAGTCGATGGGCAGACATATCGCTTGCCCAACCGTGTCCAACAACAGGACTCGC
CGGTCATATTCTTGGTACGTCAGGGACAGACAGTGTCTTTCTCCGGCAAACTCACCAGCGACCTGCCAGAAATCGAGTCG
TTCTACATTATCAAGCAGGCCCCTCTCGTTCCCTTCGGATCGGAGCAGCAACAATGAAGTCGAACAGAGGCTTCACTCTC
ATCGAGTTGATGATCGTCGTAGTAATCATCGCTATTCTTGCTGGTATCGCCTACCCCAGCTACGACGAATACGTGAAGCG
CGGGAATCGCACCGAAGGACAGGCATTACTCAGCGAAGCAGCCGCTACTCAAGAGCGCTATTTTTCACAGAACAATACTT
ATATCACTACCCAAGCCGACATCGGCAAGCTGCATATGCGCAACACATCGGGCACCACAGTGAAGTCCTCCACAGGCAAA
TACAGCCTTACCGTCGATACGGTAGCCAACGACGGAGGTTATCGCCTTATCGCTAACCAGGCATTCAACGATCTTGATTG
TGGCAACCTGACCTTGACCGCCAACGGCGAGAAAGGCCGGACTGGAAGCAAGAAGAGCGTTGCAGAATGCTGGCGCTAA

>ORF40329 (SEQ ID NO:420)
CGGACGACAAACCTGGCGGGTCGTCGAGGAGAAATGAACATGAACCCCTTACGTCTTCTCGCCACAGCTCTTGCAGCTCT
AGCTCTGGCTTGCCCAACCTTTGCCTTGAGTGCCACGAATACGTTCGAGAATGTGGGCGTGGTCGAGGATGTTCATCCTG
CCGCCGGTCTGGTAGTAGTCGATGGGCAGACATATCGCTTGCCCAACCGTGTCCAACAACAGGACTCGCCGGTCATATTC
TTGGTACGTCAGGGACAGACAGTGTCTTTCTCCGGCAAACTCACCAGCGACCTGCCAGAAATCGAGTCGTTCTACATTAT
CAAGCAGGCCCCTCTCGTTCCCTTCGGATCGGAGCAGCAACAATGA

>ORF40709c (SEQ ID NO:422)
AGCCTCTGTTCGACTTCATTGTTGCTGCTCCGATCCGAAGGGAACGAGAGGGGCCTGCTTGATAATGTAGAACGACTCGA
TTTCTGGCAGGTCGCTGGTGAGTTTGCCGGAGAAAGACACTGTCTGTCCCTGACGTACCAAGAATATGACCGGCGAGTCC
TGTTGTTGGACACGGTTGGGCAAGCGATATGTCTGCCCATCGACTACTACCAGACCGGCGGCAGGATGAACATCCTCGAC
CACGCCCACATTCTCGAACGTATTCGTGGCACTCAAGGCAAAGGTTGGGCAAGCCAGAGCTAG

>ORF40507 (SEQ ID NO:424)
TCGATGGGCAGACATATCGCTTGCCCAACCGTGTCCAACAACAGGACTCGCCGGTCATATTCTTGGTACGTCAGGGACAG
ACAGTGTCTTTCTCCGGCAAACTCACCAGCGACCTGCCAGAAATCGAGTCGTTCTACATTATCAAGCAGGCCCCTCTCGT
TCCCTTCGGATCGGAGCAGCAACAATGAAGTCGAACAGAGGCTTCACTCTCATCGAGTTGATGATCGTCGTAGTAATCAT
CGCTATTCTTGCTGGTATCGCCTACCCCAGCTACGACGAATACGTGAAGCGCGGGAATCGCACCGAAGGACAGGCATTAC
TCAGCGAAGCAGCCGCTACTCAAGAGCGCTATTTTTCACAGAACAATACTTATATCACTACCCAAGCCGACATCGGCAAG
CTGCATATGCGCAACACATCGGGCACCACAGTGAAGTCCTCCACAGGCAAATACAGCCTTACCGTCGATACGGTAGCCAA
CGACGGAGGTTATCGCCTTATCGCTAA

>ORF41275c (SEQ ID NO:426)
GTGGGGGGCGTCGGAAGAGCAGGAACTGGAGGGACGGGAGGAGAACATTACCTTCTCGATGCCCAAGGAACTGCGGGTCA
AGGCTTTGTAATCGGAATTTTTGCGCACCTGAAAAAGCCCGGCTTATGCCGGGCTTTGCCTTTTTCTTGTCTCGGCGCTT
TAGCGCCAGCATTCTGCAACGCTCTTCTTGCTTCCAGTCCGGCCTTTCTCGCCGTTGGCGGTCAAGGTCAGGTTGCCACA
ATCAAGATCGTTGAATGCCTGGTTAGCGATAAGGCGATAACCTCCGTCGTTGGCTACCGTATCGACGGTAAGGCTGTATT
TGCCTGTGGAGGACTTCACTGTGGTGCCCGATGTGTTGCGCATATGCAGCTTGCCGATGTCGGCTTGGGTAGTGATATAA
GTATTGTTCTGTGA

>ORF42234c (SEQ ID NO:428)
TCGACGTCCAGCCGGCCTGAACCGTCGGTCGCTGCGCCCTTCCCAAGCGGGGAGGGCGGTAGCAAGGTTCATTCGTCCAA
TCACCGCGTCGCCCACGAGACCGCCATGCAAATCAAACTCGCCAATCCCCGCGGCTTCTGCGCCGGCGTGGATCGCGCCA
TCGAGATCGTCAACCGTGCCCTCGATGTCTTCGGCCCGCCGATCTACGTGCGTCACGAGGTGGTGCACAACAAGTTCGTC
GTGGACAACCTGCGCCAGCGCGGCGCCATCTTCGTCGAGGAACTCGATCAGGTGCCGGACAACGTCATCGTCATCTTCAG
CGCCCACGGCGTTTCCCAGGCGGTCCGCAAGGAAGCCGAGGGGCGCGGCCTGAAGGTTTTCGACGCGACCTGCCCGCTGG
TGACCAAGGTGCACATGGAAGTGGTGCGCTACAGCCGCGACGGCCACGAATGCGTGCTGATCGGGCATGAAGGCCACCCC
GAGGTGGAAGGCACCATGGGCCAGTACGATGCCAGCAACGGCGGTGCCATCTACCTGGTGGAGGACGAGGCCGACGTCGC
CGCGCTGGAGGTGCGCAAGCCCGAAGCCCTGCACTACGTGACCCAGACCACCCTGTCGATGGACGACACCTCGAAGGTCA
TCGATGCCCTGCGCGCCAAGTTCCCGCAGATCCAGGGGCCGCGCAAGAACGACATCTGCTATGCCACCCAGAACCGCCAG
GATGCCGTGAAGGAACTGGCCGACCAGTGCGACATGGTCCTGGTGGTGGGCAGCCCCAACAGTTCCAACTCCAACCGCCT
GCGCGAACTCGCCGAGCGCATGGGCACGCCGGCCTACCTGATCGACGGCGCCGAGGACATGCAACGCGGCTGGTTCGACG
GTGTGCGTCGCATCGGAATCACCGCAGGCGCCTCCGCGCCGGAAGTGCTGGTGCGCGGAGTGATCGCCCAGCTACGTGAG
TGGGGGGGCGTCGGAAGAGCAGGAACTGGAGGGACGGGAGGAGAACATTACCTTCTCGATGCCCAAGGAACTGCGGGTCAA
GGCTTTGTAA

Fig. 3-37

>ORF41764c (SEQ ID NO:430)
AGGCCACCCCGAGGTGGAAGGCACCATGGGCCAGTACGATGCCAGCAACGGCGGTGCCATCTACCTGGTGGAGGACGAGG
CCGACGTCGCCGCGCTGGAGGTGCGCAAGCCCGAAGCCCTGCACTACGTGACCCAGACCACCCTGTCGATGGACGACACC
TCGAAGGTCATCGATGCCCTGCGCGCCAAGTTCCCGCAGATCCAGGGGCCGCGCAAGAACGACATCTGCTATGCCACCCA
GAACCGCCAGGATGCCGTGAAGGAACTGGCCGACCAGTGCGACATGGTCCTGGTGGTGGGCAGCCCCAACAGTTCCAACT
CCAACCGCCTGCGCGAACTCGCCGAGCGCATGGGCACGCCGGCCTACCTGATCGACGGCGCCGAGGACATGCAACGCGGC
TGGTTCGACGGTGTGCGTCGCATCGGAATCACCGCAGGCGCCTCCGCGCCGGAAGTGCTGGTGCGCGGAGTGATCGCCCA
GCTACGTGA

>ORF41284 (SEQ ID NO:432)
CTGGGCGATCACTCCGCGCACCAGCACTTCCGGCGCGGAGGCGCCTGCGGTGATTCCGATGCGACGCACACCGTCGAACC
AGCCGCGTTGCATGTCCTCGGCGCCGTCGATCAGGTAGGCCGGCGTGCCCATGCGCTCGGCGAGTTCGCGCAGGCGGTTG
GAGTTGGAACTGTTGGGGCTGCCCACCACCAGGACCATGTCGCACTGGTCGGCCAGTTCCTTCACGGCATCCTGGCGGTT
CTGGGTGGCATAGCAGATGTCGTTCTTGCGCGGCCCCTGGATCTGCGGGAACTTGGCGCGCAGGGCATCGATGACCTTCG
AGGTGTCGTCCATCGACAGGGTGGTCTGGGTCACGTAGTGCAGGGCTTCGGGCTTGCGCACCTCCAGCGCGGCGACGTCG
GCCTCGTCCTCCACCAGGTAGATGGCACCGCCGTTGCTGGCATCGTACTGGCCCATGGTGCCTTCCACCTCGGGGTGGCC
TTCATGCCCGATCAGCACGCATTCGTGGCCGTCGCGGCTGTAGCGCACCACTTCCATGTGCACCTTGGTCACCAGCGGGC
AGGTCGCGTCGAAAACCTTCAGGCCGCGCCCCTCGGCTTCCTTGCGGACCGCCTGGGAAACGCCGTGGGCGCTGAAGATG
ACGATGACGTTGTCCGGCACCTGATCGAGTTCCTCGACGAAGATGGCGCCGCGCTGGCGCAGGTTGTCCACGACGAACTT
GTTGTGCACCACCTCGTGACGCACGTAGATCGGCGGGCCGAAGACATCGAGGGCACGGTTGACGATCTCGATGGCGCGAT
CCACGCCGGCGCAGAAGCCGCGGGGATTGGCGAGTTTGATTTGCATGGCGGTCTCGTGGGCGACGCGGTGATTGGACGAA
TGAACCTTGCTACCGCCCTCCCCGCTTGGGAAGGGCGCAGCGACCGACGGTTCAGGCCGGCTGGACGTCGA

>ORF41598 (SEQ ID NO:434)
CCTTCGAGGTGTCGTCCATCGACAGGGTGGTCTGGGTCACGTAGTGCAGGGCTTCGGGCTTGCGCACCTCCAGCGCGGCG
ACGTCGGCCTCGTCCTCCACCAGGTAGATGGCACCGCCGTTGCTGGCATCGTACTGGCCCATGGTGCCTTCCACCTCGGG
GTGGCCTTCATGCCCGATCAGCACGCATTCGTGGCCGTCGCGGCTGTAGCGCACCACTTCCATGTGCACCTTGGTCACCA
GCGGGCAGGTCGCGTCGAAAACCTTCAGGCCGCGCCCCTCGGCTTCCTTGCGGACCGCCTGGGAAACGCCGTGGGCGCTG
A

>ORF42172c (SEQ ID NO:436)
CAAGGTTCATTCGTCCAATCACCGCGTCGCCCACGAGACCGCCATGCAAATCAAACTCGCCAATCCCCGCGGCTTCTGCG
CCGGCGTGGATCGCGCCATCGAGATCGTCAACCGTGCCCTCGATGTCTTCGGCCCGCCGATCTACGTGCGTCACGAGGTG
GTGCACAACAAGTTCGTCGTGGACAACCTGCGCCAGCGCGGCGCCATCTTCGTCGAGGAACTCGATCAGGTGCCGGACAA
CGTCATCGTCATCTTCAGCGCCCACGGCGTTTCCCAGGCGGTCCGCAAGGAAGCCGAGGGGCGCGGCCTGAAGGTTTTCG
ACGCGACCTGCCCGCTGGTGACCAAGGTGCACATGGAAGTGGTGCGCTACAGCCGCGACGGCCACGAATGCGTGCTGATC
GGGCATGA

>ORF42233c (SEQ ID NO:151)
CGACGTCCAGCCGGCCTGAACCGTCGGTCGCTGCGCCCTTCCCAAGCGGGGAGGGCGGTAGCAAGGTTCATTCGTCCAAT
CACCGCGTCGCCCACGAGACCGCCATGCAAATCAAACTCGCCAATCCCCGCGGCTTCTGCGCCGGCGTGGATCGCGCCAT
CGAGATCGTCAACCGTGCCCTCGATGTCTTCGGCCCGCCGATCTACGTGCGTCACGAGGTGGTGCACAACAAGTTCGTCG
TGGACAACCTGCGCCAGCGCGGCGCCATCTTCGTCGAGGAACTCGATCAGGTGCCGGACAACGTCATCGTCATCTTCAGC
GCCCACGGCGTTTCCCAGGCGGTCCGCAAGGAAGCCGAGGGGCGCGGCCTGA

Fig. 3-38

>ORF42233c (SEQ ID NO:151)
CGACGTCCAGCCGGCCTGAACCGTCGGTCGCTGCGCCCTTCCCAAGCGGGGAGGGCGGTAGCAAGGTTCATTCGTCCAAT
CACCGCGTCGCCCACGAGACCGCCATGCAAATCAAACTCGCCAATCCCCGCGGCTTCTGCGCCGGCGTGGATCGCGCCAT
CGAGATCGTCAACCGTGCCCTCGATGTCTTCGGCCCGCCGATCTACGTGCGTCACGAGGTGGTGCACAACAAGTTCGTCG
TGGACAACCTGCGCCAGCGCGGCGCCATCTTCGTCGAGGAACTCGATCAGGTGCCGGACAACGTCATCGTCATCTTCAGC
GCCCACGGCGTTTCCCAGGCGGTCCGCAAGGAAGCCGAGGGGCGCGGCCTGA

Fig. 3-39

>ORF2 (SEQ ID NO:3)
SPIQCQGVPGQSEPTHGCRGRHCQAPGRRREQHQYRLQRQRHQLRDDRNQQQLGPQQHPLRRRQRHPAVDEQVVRGGLRR
RLRAARCAGRSASRSATGDRL*

>ORF3 (SEQ ID NO:5)
RRSNAKEYLGNQSLLTAAGAGIAKLLDADENNTSTVFSGNGTSFGTTGTNSNSALNSILSGGVSDIRQWMNKLYGEAFAA
VYVQPGARVAVHLDQQLAIDYELKGRKVDYSSGAAHATADLD*

>ORF602c (SEQ ID NO:7)
SAWSFAEASCCGSIGRRSVCLASRSSRPRLLPIELVAPRSQTSSMLASPWGSISSLLVEHAARVSAQARPAQRRRRGLVQ
VCCCMSGSRAVIDLAALEFIVDRQLLIEMHCDPRTWLHVDGGEGLPVQLVHPLPDVADAAGEDAVEGRVAVGSGRPEAGA
VAAEDGTGVVLVGVQELGNAGPGSRE*

>ORF214 (SEQ ID NO:9)
TSCTGRPSPPSTCSQVRGSQCISISNWRSTMNSRAARSITALEPLMQQQTWTNPLLRLCAGLACALTLAACSTSKEEMLP
HGEANMLDVWERGATSSIGNSRGRLLLDARQTLRRPIDPQQDASANDQADYTRTASNEIHSQFKRLPNPDLVMYVFPHLA
GSDPAPVPGYTTVFPFYQRVQYAMPGERTEDY*

>ORF1242c (SEQ ID NO:11)
SRPGRRTGQSRVRFRARRRSSAGLLSMRPGRSASNWDRGPRCPRAPVRRMRRANAHPPGASLARRAGTQPRAAGLRTMGR
DRRGVTLRPAWRHSCSRCWAEEYPWRPVAPDSAQSLLPRPLRPALLNLRERLPVPTEAVCDRAEGFEKSPSIVLRAFARH
GVLDSLVEGEHGGVARYRGGIAAGQVREHIHHQVGIGQSFELTVDLVAGRAGVVGLVIRGGILLRIDWAPQRLPGIEEQP
ATAVAYRAGRTSLPDVEHVGLAVGQHLLLAGGARCQGQRAGQAGAETKKGVSPSLLLHERLQSCNRPCGP*

>ORF594 (SEQ ID NO:13)
PGRLHPHGQQRDPQSVQTTAQSRPGDVCVPAPGRQRSRPGTGLHHRVPLLPASPVRHAGRTHGGLLMGFFQTLLRGRTQP
QSVPADAPEDSGALDVAAAEEATERYLARLAAMGIPLPNTGSKNGATQAEASRLYDHDPSFVDLLPWAEYLPDEQVMLLE
DGRSRAAFFELVPLGTEGRDPNWMQNARDALKEALQNSFDEHETSPWIVQFYAQDEISWDNFQEQLRQYVHPRARGSAFS
EMYLALMKHHLEGISKPGGLFVDTAVSKLPWRGQQRRVRMVVYRRIRKEDAQIRGQDPAAYLKSICERIQGGLANAGIVA
SRMGGQEIRNWLIRWFNPHPDHLGQAEADLRRFYELVCRPDEPILQDELPLADGTDFSQNLFYRQPVSDATQGVWLFDAM
PHRVIVVDQLNKAPLTGHFTGETLKGDGLNALFDRMPEDTLLCITMVVTPQDMLEGHLQQLSKKAVGDTQASIHTREDVA
TVRRLIGREHKLYRGAIALFVRGRDHTQLEERCITLSNVLLGAGLVPVEPQNEVGPLNSYLRWLPSNFDPNEKRALEWYT
QMMFAQHIANLSPIWGRTTGTGHPGFTLFNRGGAPLTFDPFNKLDRQMNAHGFIFGPTGSGKSASLTNLICQMLAMYLPR
MFVAEAGNSFGLLADLAKRFGLSVHRVRLAPGSGVSLAPFADAIKLVESPDQVKVLDAEDIEASDSVQGSKADLEDDQRD
ILGEMEIVARLMITGGEEKEDARLTRADRSAVRQAILAAARTCAAANRTVLTQDVRDALYEASRSDSTAPERRARIAEMA
EAMQMFCMGADGEMFNREGTPWPEADLTVVDFATYAREGYAAQLGIAYISLLNTVNNIAERDQFKGRPIVKITDEGHIIT
KHPLLLPYAMKITKMWRKLGAWFWLATQNIDDIPASGAPMLNMIEWWLCLNMPPDEVEKISRFRELSPAQKSMMLSARKE
SGKFTEGVLLAKGKEYLVRVVPPSLYLALAMTENEEKNQRYNIMQATGCDELEAALQVAADLDKARGLPPFPIVFPDQPA
VECQDE*

>ORF1040 (SEQ ID NO:15)
VPARRASDAPGGWAFARRILRTGALGHRGPRSQLDAERPGRIERSPAELLRRARNLTLDCPVLRPGRDQLGQFPGAVEAV
RPSSSARIGLQRDVPGAHEASPGGHFEAGRTVRRHRRQQAALARTTAPRADGRLPPDPQGGCADSRTGPGGVPEIHLRAY
PRRPGERRHRRFAHGRTGDQELVDPLVQPAPGSPRPGRGGPTSLLRTGMPSGRTDPAG*

>ORF1640c (SEQ ID NO:17)
VRLGLAEVIRVRVEPADQPVPDLLSAHARSDDAGVRQAALDTLADGFQVRRRVLSANLRILLADPAVDDHPHAALLSSPG
QLADGGVDEQSARLRNALQVMLHERQVHLAEGRSSRSRMDVLPQLLLEIVPADLVLGVELDNPG*

Fig. 4-1

>ORF2228c (SEQ ID NO:19)
GEPAQVAVQRSDFVLRFDRHQAGAEQYVAQGDAAFLQLGMVAAAHEQSDRSAIELVLPADQASNGGHVLAGVDRGLGVTN
GLFRELLQMPFQHVLRRHDHGDAQQRVLGHSIEQGVEAIAFERLAGEVACQRRFVQLVDHNHSVRHGIEEPYALGGIGNR
LPIEQVLGEVSAVGQWQFILQDRFVRTAYQFVEAT*

>ORF2068c (SEQ ID NO:21)
SLCSRPIRRRTVATSSRVWIEAWVSPTAFFESCCRCPSSMSCGVTTMVMHSSVSSGIRSNRALRPSPLSVSPVKWPVSGA
LFNWSTTITRCGMASKSHTPWVASETGCR*

>ORF1997 (SEQ ID NO:23)
HPGLDPHPRGRGHRSTPDRPGAQALSRSDRSVRARPRPYPVGGTLHHPEQRTARRRPGAGRTAERSRTAEQLPALAPLKL
RSKREASPGVVHPDDVRSAHRQPVAHLGAHHRYRTPWLHAVQPWRRAVDLRPVQQAGPADECPRLHLRANWLRQVGVPDQ
PHLPDARHVPAADVRRGSGQQLRPAGRLSQAVWPLGPPGAPRPGLRRQPGAVRGRHQAGREPRPSEGAGRRRHRGLGLGP
GQQGRPRGRPARHPGRDGDRRPPHDYRWRREGRCAPDPCRSQRRPPGDPGGGQDLRRREPHGTDPRRARCALRGLQER*

>ORF2558c (SEQ ID NO:25)
VGQQAEAVARFRDEHPRQVHGEHLADEVGQGRRLAGASWPEDEAVGIHLPVQLVERVEGQRRAATVEQREARVSGTGGAP
PDGRQVGDVLSEHHLGVPLQGSLLVWIEV*

>ORF2929c (SEQ ID NO:27)
SASRTSWVSTVRFAAAQVLAAARIAWRTALRSARVRRASSFSSPPVIMRRATISISPRMSRWSSSRSALLPWTESEASMS
SASSTFTWSGLSTSLMASANGARLTPEPGARRTRWTERPNRLAKSASRPKLLPASATNIRGRYMASIWQMRLVRDADLPE
PVGPKMKPWAFICRSSLLNGSKVNGAPPRLNSVKPGCPVPVVRPQMGDRLAMC*

>ORF3965c (SEQ ID NO:29)
APVGPYQAVDVVAAIHPRAALSAGRYPGDRLPSVESAAAPLSVQERISLASAGHPLRGSAGSGSGCRSGSGSANSELSFV
LALHCRLVWENNGEGWQAARLVEIRCDLQGRLELVAAGGLHDVVALVLFFVFGHGQGQVETRGNHTDEVFFALGQEHALG
ELAAFLAGREHHRLLRRRQLAEPGYLLYFVGGHVQAQPPLDHVQHRRPGGWDVVDVLGGEPEPGAQFPPHLGDLHGVGQQ
QRVLGDDVPLIGDLDDWPALELVAFGDVVHGVQQRDVGDPELGGVAFARVRCEIHHGKVGLRPGRAFAIEHLAVGAHAEH
LHGFRHFGDPRAAFWRGAIAPGGLVERIAHVLGQYRAVRGGAGPGRRQDRLADGAAIGTGQARIFLLFATGNHEAGDDLH
LAQDVSLVVLEVGLAALDRVRGLDVFGVQHLHLVGALDQLDGVRERRQADAGARGEAHPVDREAKPLG*

>ORF3218 (SEQ ID NO:31)
GAHHHQAPAAAALRHEDHQDVAETGRLVLARHPEHRRHPSLRGADAEHDRVVVVPEHAPRRSREDIQVPRAVAGAEVDDA
LGPQGKRQVHRGRAPGQGQRIPRPCGSPESLPGPGHDRKRRKEPALQHHASHRLRRARGGLAGRSGSRQGARPATLPHCF
PRPTGSGVPGRMRVLNSLTQNLIDNLTQILQNPEEDALQTLRICAPVLIEELQQIQLRAVDRRDIVPQIKQLLDEWLQQH
PQPDTAQQALIEAVDRAEILQRRQA*

>ORF3568 (SEQ ID NO:33)
PKTKKRTSATTSCKPPAATSSRRPCRSQRISTRRAACHPSPLFSQTNRQWSARTNESSEFADPEPDRQPDPDPAEPRRGC
PADAKDMRSCTDRGAAADSTEGSRSPGYRPADKAALG*

>ORF4506c (SEQ ID NO:35)
VNKFVVFRTFLQSSLVQFRKVQCAARQPAPVAGRLSEDRIDSAPEGFGAALDPRALHQASLVAGRLAMHLQGKMAPNQVH
VRMAVYPALKPRGVDLAEGALQVGVFIDRPARFRIAVEAVVGWQALHQKLYPYGGCSQQDQQQPRPGQGGTLKSFGCPAA
LQESHACLRCRISARSTASMSACWAVSGCGCCCSHSSKSCFICGTISRRSTALS*

>ORF3973 (SEQ ID NO:37)
GRGPRGDPTAEASVRLLKGGWAAKRFQGPALPWAGLLLVLLAASAVGVELLVKGLPANHSLYGDAKARWTINEYADLECP
FCKVYTPRLKRWVDSHPDVNLVWRHLPLQMHGEAARHQARLVECAGIQGGAKAFWSAIDAIFAQSAGNGGGLPGGTLDFP
ELDQARLEKCAKDNELIDSDIKLDIDIARSKGITATPTLVIRDNQTGRSVKLEGMADETTLLSAIDWLAKDL*

Fig. 4-2

>ORF4271 (SEQ ID NO:39)
TWFGAIFPCRCMARRPATRLAWWSARGSKAAPKPSGALSMRSSLSRPATGAGCLAAHWTFLNWTRLDWRNVRKTTNLLTQ
ISSWTSTLHGRRALQRPRPSSSGTTRRDEA*

>ORF4698 (SEQ ID NO:41)
EIGEDSNIPLLVLQDALHFTWQNLDLLPIHNLYHSLVAGAGEAKPQLHCRPSIDVNALEQALHDFDHSLISVSQLHTGIM
LPRTCRRHPYLCTWQRSITARKNTPPTS*

>ORF5028 (SEQ ID NO:43)
FPAALSEVILSAVCTFLEPVQTHASSSLPPWPAATNAGRWRTTGTAEQRESGRNLGHHRQGSSGLCHRIVARSVSGRPGT
PRGATDCGLAPGSTACSSGV*

>ORF5080 (SEQ ID NO:45)
NRYRPMPLHHSPPGRRPPTLAVGVLLVLLSSASQAETWVITDKAHPVSATGSSRVLFLDAQEHLEEQLTAALPQDPQHAQ
AAFKRLLQSPDGRRLQAELVKAQQDVADAWSLGVEKIPAVVVDRQYVVYGEPDVSRALELIAKARRSR*

>ORF6479c (SEQ ID NO:47)
FVSVSLLEVGTADEHLPLALAAGVGTPERPGVLPVDGLRLRPRVGKHRAVEAQGWGQLLPFPGRGIALFQLARRPVAVLG
GCAHGEVDVELADSRGDIAGALGDDGCRLVVVGLVQEAAARIEVPPHVAGEDSTHLAQPWDQRFGVHLLGNSMPPANGVQ
CAEKVRHQRDGGARANVPRGAGEPAERGATRMADHIRFLEAADAVLGLVVCGRVIAGLGEWIRCTQRRYLGPGVAPGIRV
AGDDCVRHVVADLDRRLHFAAMRAAEQPVTDPDDLVFEALRGKGGGDDGSAVDRGRGREREAEGGGRRCQAAEVEAGHQR
DLLALAISSRARETSGSP*

>ORF5496 (SEQ ID NO:49)
ANRQGQEVALMTSLNLRRLAAAAATFSLSFTASAAINSAAIVSSTLSPQCLEYKVVGICYWLLCGPHGCKVKTSVKVRHY
VPDAVVSSYANTGSNPWTEVSALGTPNPLAQAGNDATTNYKAENSIGRFKEADVIGHPGGATFSRFASASGYVCPGATVP
LVPYFLSTLDAIGWRHGIPEQVYPEALVPGLREVGGIFSGDMWGNLYPRSGFLHQTDDYKTAAVIAQRAGDITTRIGQLH
VYLPMRAAPKDGYWPAGELKEGDASTGKWQELTPSLSLNCAVFPNSGPKTQAVDGEHAWALWRPYSCCQRKGQMFICSTD
FQ*

>ORF5840 (SEQ ID NO:51)
RDHKLQGREQHRPLQGSGCDRPSWWRHVQPVRQRLWVRLPWRHRPAGAVLSQHTGRHWLAAWNSRAGVPRSVGPRAARGG
WNLLRRHVGEPLSAQRLPAPDRRLQDGSRHRPARRRYHHANRPAPRLPPHARSPQGRLLAGGRAERGRCLDREMAGADPI
PEPQLRGVSQLWAEDASRRRGARLGALASLLLLPAQGADVHLQYRLPIRTRRRIMRMNITSVALMWLLAAQLAQADDPIN
VSKTGTVLSDEVLYSIGGGSAVSMGSAGQMDSIGVGFGWNNDMMCGNMNLSTTLENQLNGATQGFQNIMGSVIQNATGAV
MSLPALIIQRANPQLYNLITNGILQARIDYDRSKGTCKTIAEKMADIAGEQTGWGKIAEGQALGATLASDGKDAVSALEA
VEKKGGNDGVTWVGGDKAGGSGQKPIRIVNDVTRAGYNLLTSRSVNDSSSVPSATCNNGLVCNTWSSPQEAAAFATRVLG
EQQQQTCEGCQKTVTAAGVGLTPLIQETYDKKLQSLQELLSKSKPLTAENLAAAGTDALPITRGVIEALRDERDQDVLAR
RLASDVSLMDVLSKALLLQRLMFAGAKEPNVAANGLATQAVDQQTSLLQQEISNLKTELELRRELASNSPMRVIERGQQR
ASGSSGVFESAPDADRLDRLQAPSAAGGKSGGRP*

>ORF5899 (SEQ ID NO:53)
SAILVAPRSAGSPAPLGTFALAPPSRWCRTFSAHWTPLAGGMEFPSRCTPKRWSQGCARWVESSPATCGGTSIRAAASCT
RPTTTRRQPSSPSAPAISPRESASSTSTSPCAQPPRTATGRRAS*

>ORF6325 (SEQ ID NO:55)
ASTARCFPTLGRRRKPSTGSTPGRSGVPTPAASARGRCSSAVPTSNKDTETNHANEHHLGRANVAARSATCPGRRPDQRV
QDRHGAQRRGPLQHWRRQCGEHGQRRPDGLDRRRLRLEQRHDVRKHEPEHHPGEPAQRCHTGFPEHHGLSHPERDRRGHV
AAGVDHPAREPSALQPDHQWHPAGADRLRPLERDLQNDRRKDG*

Fig. 4-3

>ORF7567c (SEQ ID NO:57)
QCLAEHVHQGDIGRQAARQDVLVTLVAQRLDDAAGNWQSIGAGRSQVLCSQWFALRQQLLQRLELLVVGLLDQRGEADAS
SRHRLLAAFAGLLLLLPQYPGGECGGLLGGGPSVADQAVVASGGRHARRIIHRAAGQQVVARPGHVVDDANGLLAGAAGL
VSTNPGYAIVAAFLLHCFEGGYGVFPVRGQCGAQGLAFGDFPPAGLLASDVSHLFGDRFASPFRAVVVDPRLQDAIGDQV
VELRVRALDDQRRQRHDRAGRVLDD*

>ORF7180 (SEQ ID NO:59)
FVERAFRHLQQRPGLQHLVLPPGGRRIRHPGTGGATATDLRRLPEDGDGCWRRPHPADPGDLRQEAPVAAGAAVEEQTTD
CREPGCGRHRCSANYPRRHRGAARRA*

>ORF7501 (SEQ ID NO:61)
PGRPGAPPGVRCLPDGRAQQGTATAAPDVRRRQGAQRRRQRPGHPSRRSADQPPAAGDLQSQDRTGTPSRVGQQLPHAGH
RARATTRLRVQWRVRVGARCRSPRSPAGPLCRRRQVGRETVMADTLTTRKLLGQLLVGVLIVIGLAVVGTLLSLFALNHF
GGIQGLEAWRQSNYWSLFAWRALLYCALAIAWFRQRKELSAHERQRIRRIEILVLLLVLLIEFSKAYFRTGGAA*

>ORF7584 (SEQ ID NO:63)
CSPAPRSPTSPPTAWPPKPSISRPASCSRRSPISRPNWNSVASWPATPPCGSSSAGNNAPQGPVACSSRRPMPIASIACR
PPLPPAASREGDRDGRYAHHPKASRSATGRSADRHRTGSGRYAAQSLRPEPLRWHPGPGGLAAKQLLELVRLAGAAVLRP
GHRLVPAAQGTERA*

>ORF8208c (SEQ ID NO:65)
RSCCASRAEVGFAEFDEQDQQQHQDLDPPNALPLMRAQFLALPEPGDGQGAVQQRPPGEQAPVVALPPGLQALDATEVVQ
GEETEQRTDHCQSDDDQHSDQ*

>ORF8109 (SEQ ID NO:67)
AAAHSADRDPGAVAGPAHRIQQSLLPHGRRSMTFMTNDYLEYYLTLLGWIINNGIWNMISDTGLFAVPFAAIVMREWLKV
RGEGADEGNKGVLSLARIETHIYVGYIVVALAGIPVVNVSFDTIEFDQTRAQQCQYNLPAPADTGWSSSFSSLAGKSAQM
PLWWAMMHALSKGFTSGAIAAIPCGTDLRQMRMEVDNTRVNNPLLAQEIADFSRDCYGPSRARLFMRQPDLGSVAEDNKA
LQDLNWIGSRFLLNTPGYYDTDYSKSPRQSWPYNATRDAGLPQVGGGGGYPTCKQWWADSGIGLRDRIKDQVDPDLMTSF
LKWAKWLNQDEVTEAVIRQVISPSSQVKGNVYTDYGGQVGGTVWNGIARTAGTFGVAVGSLAYFPAMDMVRQALPMVMSF
LKMAMVICIPMVLVIGTYQLKVAMTMTVVFFAMMFVDFWFQLARYIDSTILDAFYGSGSPHLSFNPVMGLNTATQDAILN
FVMGSMFIVLPLLWMTAIGWSGIQAGSVLNGLSRGTEGVQAAGKEAGNRVKNAV*

>ORF9005c (SEQ ID NO:69)
VSPPLLAGWVAATTAHLRQAGIAGGVVGPRLTGTLRVVGVVVPRGVQQESGADPVQVLQRLVVLGDGAQVGLPHEQPRTG
RPVAVSGKISDFLCQQRIVHARVVHFHSHLPQIRAARNGRDGAAGEALGQGVHHRPPERHLRTLAGQAAEGARPAGVRRC
RQIVLALLGASLVELDGVEAHVDDRDPRQGDHDVADVDMRLDAGERQHSLVALVGAFPTNFQPFAHHDGRERHREQASIR
DHVPDPVVDDPAEEGEVILQVVIGHEGHAAPPVRK*

>ORF8222 (SEQ ID NO:71)
LPGVLPHPPRLDHQQRDLEHDLGYWPVRGAVRGHRDARMAESSWGRRRRGQQGSAVSRPHRDAYLRRLHRGRPGGDPGRQ
RELRHHRVRPDSRPAVPIQSAGTGGHRLVELLQQPGRQECADAALVGDDARPVQGLHQRRHRGHSVRHGSAADANGSGQH
AREQSAAGTRNR*

>ORF8755c (SEQ ID NO:73)
QSLEKSAISCASSGLFTRVLSTSIRICRRSVPHGMAAMAPLVKPLDRACIIAHQSGICALLPARLLKELDQPVSAGAGRL
YWHCWARVWSNSMVSKLTLTTGIPARATTM*

>ORF9431c (SEQ ID NO:75)
LKPEVDEHHRKEDDRHRHGNFQLIGADDQDHRNADDHCHLQERHHHRQCLADHIHRREVCQAAHRNAEGSCGSRDAVPHG
AAHLPAVIGVDVTLDLAGG*

Fig. 4-4

>ORF9158 (SEQ ID NO:77)
RLHRLRRAGGRHRVERHRENRRNLRRCGGQLGILPGDGYGPPGTADGDVVPEDGNGHLHSDGPGHRHLSTESCHDDDGRL
LCDDVRRLLVSVSQIYRQHDT*

>ORF10125c (SEQ ID NO:79)
VIAGCLPLGARRLMMNAHTNKGFASRIGFGLGMLVRFCLHDRRPALRWVKRVSLFLLVALVVSQNFMWLAGVSMTLLCVF
LVGFALVKGDISVSKGSPSRDVSTMTSQAETESVAELFDYQAAHHYRD*

>ORF9770 (SEQ ID NO:81)
SNSSATDSVSACEVIVETSRLGDPLETEMSPLTKANPTRKTHSRVIDTPASHIKFCDTTRATNKNRLTRLTQRRAGRRSC
RQKRTSIPRPKPIREAKPLLVCAFIINLLAPKGRHPAITYTPKKMIWQALWHIMPLAICRLEYLMATRNVVLPDPLEQDI
NELVETGRYQNRSEVIRAGLRLLLQQEAQIAKLETLRNATSSGLMQLERGEYDEITSDELAQYLDELGNQASH*

>ORF9991 (SEQ ID NO:83)
SWTAIMQAETHKHTQTKTDPGGKAFVGVRVHHQSPGSQREASCYHLYAEKDDLASIMAYYATSYLPTGVPHGNAKRRPSR
SAGAGYQRAGGDRPLSESQRSHPGRLAPAAATGSPDRQARNPPQRNIQWADATGARRVRRDHQRRTGPIPRRARQPGEPL
KHGQVPHLS*

>ORF10765c (SEQ ID NO:85)
HLVCRHPVEDEVPGPNNLTDIGHRVAVNEVDAAQASSQFFTADAAYLLWVCRNCFQRRPYECLVTSARGIAEVVVGEAQD
IDDVRLGIMRDAVLGHASVARLVAELVEVLGQFVAGDLVVLAALQLHQPTGCCVAEGFELGYLGFLLQQQAQACPDDFAA
ILIAAGLHQLVDILLQRIGKDDVSRCHEVLQSADS*

>ORF10475 (SEQ ID NO:87)
SMAKYRISHDAQADIVDILRFTHNHFGDAARRRYQALIGAALEAVATDPQQVGSISREELGAGLRSIHLVYCHSMPNVGK
VVRPRHFVFYRVATDQVLEVVRVLHDAMDVDQHLPQR*

>ORF11095c (SEQ ID NO:89)
SRMQAVVSTNANAWSGGMQSSGQATAIAHQPWGTCWWMFTRSLFVFAGANAALSAFRQALSGRAFTLVNHSLRPSSPFPL
WAICSCYSCSSLGQVLIHIHGVVKHANHL*

>ORF11264 (SEQ ID NO:91)
TAVRRDLLKLMGCTHIEADYIGGLRCSTAPEGTWVAHGFHGPIVDVIDDSAGFFSTHRLALHYPAQCGLAVDQAIPRTAI
HVASPLMHVCIGKVVVISAWMC*

Fig. 4-5

>ORF11738 (SEQ ID NO:93)
EEVIMKLQAYRLQNYRRLRDVVIELDDEISIFVGANNSGKTSAVQGLYSMLRGEVKKFELFDFSAALWAEIDAVGRTPPG
DEDAPKRLPSILLDLWFRVGEDDLATAMSLLPSTEWDGKCVGIRVAFEPRDAHELVWKFHELHEKANNAAVALAAKRKAA
GEQAVEAGAEDAAAVVADAGEYKPWPESLTKYLTKELSKEYTFRYYVLDERAFVGYQAREADYEPLPLGKEPGGAAILKS
LVRVDFLRAQRHLDDPDAGSSDRAESLSRRLSRFYHRNLEKRGDDHAALKALDTSEKELNFHLKEVFNDTLTRLAKLGYP
GVNNPEIVIRAALDPTTVLGQDAKVHYVIPGVASAQLPDSYNGLGFKNLVYMVVELLDLHEQWKAEDDKRAPLHLVFIEE
PEAHLHAQIQQVFIRNVLRLLEDANDHATLFHTQLVITTHSPHILYERGFSPIRYFRRVNDQLGHHTDVRNLSLFKTGAS
DAPAREFLQRYLKLTHCDLFFSDAVILVEGNVERLLLPAMIELVAKRLRSSALTILEVGGAFAHRFQELIAFVGLTTLVI
TDLDSVTVKTDAEKAAAQGAGAEGAVDGDDEDEDDDLKPFELEDDDEAEPSGKKKSKKRGSTCHAHVEGAVTSNQTLISW
IPKKRSMAELWEVTAEQKTLSLAEDSSAGVRVAYQTKVSVTVGATTSQLCGRTLEEAFGLENADWCQAEANRSVGLKLKR
APSSPEELAEKLHDRVVGKNFDKTRFALEVLASGPLNGWKVPAYIAEGLAWLEAKVAHELEADAAIATEVATIEPTTADV
VAIIVDPGQTA*

>ORF12348c (SEQ ID NO:95)
RKVYSLLSSFVRYFVRLSGQGLYSPASATTAAASSAPASTACSPAALRLAASATAALLAFSCSSWNFQTSSWASRGSNAT
RIPTHLPSHSVLGSSDIAVARSSSPTRNQRSKSMDGNLLGASSSPGGVLPTASISAHSAALKSKSSNFFTSPRSIEYRPW
TADVFPLLLAPTKIEISSSSSMTTSRSRR*

>ORF12314c (SEQ ID NO:97)
GTSSGFLAKACTRRHRPPPQPRLPRPPPQLAPRRPCAWPQARQLRCWPSHVVHGTSRRARGHPEARTLPGSRRTCRPTQC
SAAATSQWRGRLHRRGTRDPRVWTVTFWAHPHRQGASCRPHRSRPTAPH*

>ORF13156c (SEQ ID NO:99)
RQIAHIRVMAQLVVDAAEVPNGRESAFIEDVRGVRGDDELRVEQSRVIVSILKEAQNVPDEDLLDLRVQMRLRLLNEDQM
KRSSLVILGFPLLVQVEQLNHHVDQILEPQAIVAVWQLGGSYARDHVVNLGVLPQDSGRIQGRPNHDLRIVDARIAELGQ
AREGVIEDFLQVEVQLLLRGI*

>ORF12795 (SEQ ID NO:101)
LPPNCQTATMAWGSRIWSTWWLSCSTCTSSGKPRMTSELRFIWSSLRSLRRICTRRSSRSSSGTFCASLRMLTITRLCST
RSSSSPRTPRTSSMNADSRPFGTSAASTTSWAITRMCAICRYSKRARPTLQRANSCSGI*

>ORF13755c (SEQ ID NO:211)
ATRTPALESSASDSVFCSAVTSQSSAIDRFFGIQLMRVWLDVTAPSTCAWQVLPRFLDFFLPLGSASSSSSSSKGFRSSS
SSSSSPSTAPSAPAPCAAAFSASVLTVTLSRSVMTSVVSPTKAISSWNRCANAPPTSRMVRAEERRRLATNSIIAGRSRR
STLPSTNITASEKKRSQCVSFRYRCRNSRAGASDAPVLNSDRLRTSV*

>ORF13795c (SEQ ID NO:213)
CRRTHRHRNLGLVSYPNPSAGILSQRQRLLLRRDFPELCHRPLLRDPADEGLVGRDGTFHVCMAGAATLLGLLLATRFCF
VVVFKLEGLQVVVLVLVISVNGALSACALRGGLLGVRLDRHAVQIRDDQCCEPNEGDQLLEPMRERTTDFKDG*

>ORF14727c (SEQ ID NO:215)
QEVGELKDVLVAKYALGVVTAHAVVERPDAGHSLQASDISLLVGLVPVARGLPDTRAVILELFFKFGDPPADVVLQPNLD
VGLERLCNRPVKAVDGRDMNQRVIVDVRQDFIGGKVGIRDARDYLLPRPHAGAVLRDHPVQRLDEGGGLAGTCASTNHEG
LRRRRYNACVDLAVGVGIWAINSSAHAVCPGSTMIATTSAVVGSIVATSVAMAASASSSWATLASSQAKPSAMYAGTFQP
LSGPLASTSSAKRVLSKFLPTTLSCNFSASSSGLLGARLSLRPTDRFASAWHQSAFSRPKASSSVRPQSCDVVAPTVTET
LVW*

>ORF13779 (SEQ ID NO:217)
RWVRRHHSSAAAHLRRPLVLRTRTGARLRQTGRSASSSSAHRAALKSWLRSYTIGWSARTSTRPALRWRYSQAGRSMAGR
FPRTSPRAWPGSKPKWPTSLRRMLPSPPRSRLLSRLQPMLSLSLLTRGRRHEQTN*

Fig. 4-6

>ORF14293c (SEQ ID NO:219)
GRWSCRNLRQHEPRRAAEASLQCMRGSRGRRRYLGYQFVCSCRLPRVNNDSDNIGCSRLNSRDLGGDGSIRLKLVGHFGF
EPGQALGDVRGNLPAIERPACEYLQRKAGLVEVLADHPIV*

>ORF14155 (SEQ ID NO:221)
PGADGMSRRIDSPDTDADREIHACIVATPPQPFVVRAGAGSGKTTSLIKALDWVISEHGASMRARKQIVACITYTDLATN
EILADVNDDPLVHVSTIHSFYWSIAKTFQADIKVWLQNDIRRRISELEEEFENYSSRVRQTTRDRNKADQERYVRSLEAV
AGVRTFNYGVGSDYAKGILGHEDILQLADFLLQNRPLFRRVVALSYPFVFIDESQDTFPGVVKSFKEVEAQMQGKFCLGF
FGDPMQSIFMRGAGDIQLEDHWRAITKPENFRCAKQILDVANAVRAQGDGMEQVRGLHERVDGNLKLVEGSARMFVLPNT
LNRTEALARVRAWSSATNNDEGWTTPDIAVKILVIVHRMAANRLGFGGIYSALNDKTSDAMKQGMQDGTGWPVRPFLSFA
LPIVAAVKAGNEFAAMSLLREFSPRLAPAALTGRRAADVLRELHAAASRLVAMLDEAGTTIGDIALHLCDTGLFEFDERY
ARVLGFVRDIADTAQEPEAADAVPAEGLSLDATMAKFFNCSAQELWPYERYVSEGSPYATQHGVKGAQFERVMVVMDEEE
SDYRTYNYERVFASAEARAADRARALDGDENTWSRTLRLLYVCCTRAQRGLVLAFFVADPATTLENVVASGILPRSAVFT
QEVLVGWP*

>ORF14360 (SEQ ID NO:223)
SRASRIPTLPPMKSWRTSTMTRWFMSRPSTAFTGLLQRRSRPTSRFGCRTTSAGGSPNLKKSSRITARVSGRPRATGTRP
TKSDMSEAWRLWPASGRSTTAWAVTTPRAYLATRTSFSSPTSCYKTARCSDGSWR*

>ORF15342c (SEQ ID NO:225)
EGSNGPTGAVLHPLLHGIRRLVVQRRVDAAEAKPVCGHAVHDDKNLDCDVWGCPTLVVVRRRAPRSDSCQSLGSVQRVRQ
DEHPGRPLHQLEVPIDPLVQPADLLHAIALRAHGIGDVKDLLGAAKVLRLRDGPPMILKLDVPCASHEDRLHRVAEKTKA
ELALHLGFHFLERLHYTRERVLTLIDKHERVAQRHDPSEQRAVL*

>ORF15260c (SEQ ID NO:227)
MPPKPSRFAAMRCTMTRILTAMSGVVQPSSLFVAELHARTLAKASVRFSVFGKTNIRADPSTSLRFPSTLSCSPRTCSMP
SPCARTALATSRICLAQRKFSGFVMARQ*

>ORF14991 (SEQ ID NO:229)
RRQCRARAGRWHGASPRAAREGRWEPQAGGGVGPDVRLAEHAEPNRGFGKSPSVELGDEQRRGLDNPRHRSQDSCHRAPH
GRKPAWLRRHLLGAERQDVGCHEARDAGRHRLARSTLPKFCATDRCSCEGRQ*

>ORF15590c (SEQ ID NO:231)
RSSNSKRPVSQRWRAMSPMVVPASSSMATSLDAAACSSRNTSAARRPVRAAGARRGLNSRSRLIAANSLPAFTAATIGSA
KLRKGRTGQPVPSCIPCFMASDVLSFSAE*

>ORF15675c (SEQ ID NO:233)
SFGRNCISSLGLLSGVSNIPDKPKNTRIALVELKKTRVTEMESYVTNGGPCLVQHGDKPRRSSVQLSQYIRGTSAGQSRR
RQARAEFPEQAHRRELIAGLHSCNDR*

>ORF16405 (SEQ ID NO:235)
IDSLRKCVGSLEKCCFACKEIIHVHAIRCRQCGESQGWRRFMSSPTSVVALVLSLLSIAATKPVERLFDAQRAELQISIT
GGDYKAAQLMLTNNGSKPATLVSFEITSKATTNTKTWFLVSNTDGEILEPGKTYKIRASTDESIPKIVEAERRTILKSQY
ALADNCELTAKYIEATGQKVVRVQPFMCDTPPEKGGLPPGKPGIPIWYLGQE*

Fig. 4-7

>ORF16925 (SEQ ID NO:237)
RPRGRRLCVCNRSCATHLLKRVACPLVNLAYPFGTLVKNDVFMPPWALTPIKQSCVRSSNTSLAQLNDCYVYGCCRYVIP
WPYAYEVNSESVQWTIFLLGVDCSGKVIYFRNTARVGPFLAASIYRPWYGSDALVLHFTK*

>ORF17793c (SEQ ID NO:239)
AKMIVIDKNLEHLVAQCAICEKTLFDEFSLKIQLGHTYYEPKSLPASASIVYGSHPAPSTFFLEPKEIQQNLVLKSGEQV
ITCSKHRYKIPLDYFGLVQTKGTLARLFVQVTCNDGQVEPGFDGYVTLEIVNMSPWTIEIPAVSDIAQLYLVKCSTSASE
PYHGRYMDAAKKGPTLAVFRK*

>ORF18548c (SEQ ID NO:241)
RTMAGWPRLAAQGRRTNLMSVLQIKGRTTKSHTDFDAASYSSNSLILTDAGDERIEEFSLELSVGEGWSDNYSGNDKNLW
RIVDGMTIRGHDSVVVEAAEEIKVPHNRYGIVLPTGSLFLSRGVLVASAKVEPAFDGKLKLRIFNTTNKNVCLTKGEKLG
SVIFFSTESTHTQSPIKRGSEISTLPITRRARLKKWFSLNPTIWVGWTLNLIGSSLVSSLIMYAVYYKVVLEHQSQPPQS
QQNAQPSPNEVKPK*

>ORF17875 (SEQ ID NO:243)
TAYIIREDTRELPIKFSVHPTHMVGLSENHFFNRARRVMGSVDISLPRLMGLWVCVDSVEKKITEPSFSPLVRQTFLLVV
LNILSLSLPSNAGSTFAEATSTPRERKRLPVGRTMPYRLCGTLISSAASTTTES*

Fig. 4-8

>ORF18479 (SEQ ID NO:245)
SVTHSSDLSFVLGLRDAATLPLSFIPADIPGYRLKDDVRKACTNLNFKRLAVIVGERERHRPYITWRQHTGTERYPASEQ
RASRKKKRRQIFRQIEFFHGARQISLARFHDEAVIRVCEHDLAGRGASRRFSQASTPYCQAREACESEVKSNAFRGGQLT
VGKVLD*

>ORF19027c (SEQ ID NO:247)
MIYSPHSLLKLVRDGKLIKHLAHRELTTPEGVGFDLRLAGLSRLTVGGGSLRESTRRTPASEVVLADPDDCFVMEPGKTY
LASTMEEFDLPEDLAALFFPRSTLFRSGITFSSSVLPPGYVGPMTFALTNNHSEAFEIQIGARFAHVIFQAVSGDIGRYK
GQWQGGRVSQPKDEGQI*

>ORF19305 (SEQ ID NO:249)
WPFSACRLFGMTGQVGCKRWSAPMQLGGHVRCNYAVEPGPVPPKQSIRPRWHIANKIPFPATVVLSLLPALIWRKSPLHE
SSWSLPCFNSFPGYPGSRPPPQQPKLPQGDSSFL*

>ORF19519 (SEQ ID NO:251)
SGGKARSMNRHGASHVSTPFLDIQEAVPHPNNQSCPRGIHPSSEQHGTARHASPPAATGEHLAARLAIQAAIRGDLPAAT
GQLCRAGPAAPCFGKSSPCPSRRDDRSRPGDRGLRTQGTADLPAPDRRSAGVTVSPG*

Fig. 4-9

>ORF19544 (SEQ ID NO:253)
IVMEPPMFQLLSWISRKPSPTPTTKAAPGGFILPLSSMELLGTPRRRQLLENIWQRASLSKQQFEEIYRRPLANYAELVQ
QLPASENHHHAHPGGMIDHGLEIVAYALKVRQTYLLPIGAAPESQSAQAEAWSAAAAYGALAHDIGKIVVDLQVELQDGS
TWHPWNGPINQPYRFKYVKSREYQLHGAASALLIHQLLPRTALDWLSRFPELWAQLIYLFAGQYEHAGILGEIIVKADQA
SVAQELGGNPDRALAAPKQSLQRQLADGLRFLVKDKFKLNQPSGPSDGWLTQDALWLVSKPAADQLRAYLLAQGIDGVPS
SNAPFFSMLQDQAVIQTNAEDKAIWTATVDNGAGWRNKFTLLKIAPALIWTDAAERPSPYSGSLVVEDGTASTEKPETTC
EIPNGPAEQQQAPETKMMLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRG
EENLQQPLGTKEPTDCAPEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQE
HPVLEKLAQAKETTGWKLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDA
EGGVE*

>ORF20008 (SEQ ID NO:255)
ATGRQHLAPLERTDQPAIPLQVREVPRIPAPRRCLSTSHPPTATAHCTRLAQSLSRAVGSIDLPVRWAVRARRDPRRDHR
EGRPGLSCTGARRQSGSSSGCTEAVAAAAVGRRPSLLGEGQVQVEST*

>ORF20623c (SEQ ID NO:257)
RELVSPSSTVVYRGRPNGLVLGICLDDGLVLEHAEERRVRGGHPIDTLGQQVGSQLIGSRLAHQPECVLGQPSIRRAARL
IQLELVLHQEAKAVCQLPLQRLLRCSQSSIRIAS*

>ORF21210c (SEQ ID NO:259)
RLKIPGVTSMAVPSTVCTKALVSLMNRRRAAMPDFIQPTNPCPRSVLLLGINTSSIASGAQSVGSLVPSGCCRFSSPRVF
VGEAAGESWLVSSSSSGELILPRSAYKSSVSS*

>ORF21493c (SEQ ID NO:261)
AAAASFHTSFGIGDDREAWVVQRLLREQQFGILEQVGLELFRLARTRNLDGPDVQVFTGLPMKPLFFERALHQLPAGRLL
GLGQFFKHRMLLDIALENSWRDQHGGPVYGMHQSLGVVDEQATGRDARFHPTNESLSQISTSARHKYIFNCFRSAICWLL
GP*

>ORF21333 (SEQ ID NO:263)
TSGPSRFLVLARRKSSRPTCSRIPNCCSLSSLWTTQASRSSPMPKEVWNDAAAAHRGVHLRARSPGSQREDLPRRDQGAA
QALRSYGNRTGRGPPGCPGMATQGTGTRPVQAELEHVLESSADDLGLCHRA*

>ORF22074c (SEQ ID NO:265)
VRHKPLRYWHYELFVSLLCLAADEDQLIFPVDVALADAQQRIQPDAGEVEDLKRAKPEPGGRDPCAFAGAALCAVLHVEP
IAGEQDGFGGDGFACSPGGNDGGLSERVDLGVRHQLMLDGIAPDRPQMIRVRVPAPLGQALFQYLASPSQDSPVVHVLYG
CRRTEVLEQRLGRGAVDLRAGFPEIVREDVLLGELLRRHSTPPSASVMTVRLGLSRGCSGNSNLGSWSR*

>ORF21421 (SEQ ID NO:267)
AASGQPKPHGHHRCRRRCGMTPQQLTEEYIFAHDLREASAKIYRAATKALLKHFGPTATVQDVDHRAVLGWRRKVLEQGL
SKRSWNTYSNHLRTIWGYAIEHELVTHSQVNPFRKTTVIPPRRASKTVAAEAILLARNWLNMQDGAERCTGERARITPAW
FWLCTFEVFYFTGIRLNALLCIRKRDIDWENQLILIRGETEKTHKEFVVPITEGLVPHLSRLLQEADRAGFADDDQLFNV
NRFSPHYKSKVMNSDQVEAMYRKLTEKVGVRMTPHRFRHTLATDLMKAPERNIHLTKCLLNHSNIQTTMSYIEADYDHMR
AVLHARSLAQGALENVRKVDYSGSPQASAKPKPCGQPLARVSEAPPPEARTEPAEPREHTPGTGIQGGPTAWEADALPQP
PDTFEPSVLFTLMAQNLSNRAASASAAPAATSGSGGWGSAARSNLA*

Fig. 4-10

>ORF20008 (SEQ ID NO:255)
ATGRQHLAPLERTDQPAIPLQVREVPRIPAPRRCLSTSHPPTATAHCTRLAQSLSRAVGSIDLPVRWAVRARRDPRRDHR
EGRPGLSCTGARRQSGSSSGCTEAVAAAAVGRRPSLLGEGQVQVEST*

>ORF20623c (SEQ ID NO:257)
RELVSPSSTVVYRGRPNGLVLGICLDDGLVLEHAEERRVRGGHPIDTLGQQVGSQLIGSRLAHQPECVLGQPSIRRAARL
IQLELVLHQEAKAVCQLPLQRLLRCSQSSIRIAS*

>ORF21210c (SEQ ID NO:259)
RLKIPGVTSMAVPSTVCTKALVSLMNRRRAAMPDFIQPTNPCPRSVLLLGINTSSIASGAQSVGSLVPSGCCRFSSPRVF
VGEAAGESWLVSSSSSGELILPRSAYKSSVSS*

>ORF21493c (SEQ ID NO:261)
AAAASFHTSFGIGDDREAWVVQRLLREQQFGILEQVGLELFRLARTRNLDGPDVQVFTGLPMKPLFFERALHQLPAGRLL
GLGQFFKHRMLLDIALENSWRDQHGGPVYGMHQSLGVVDEQATGRDARFHPTNESLSQISTSARHKYIFNCFRSAICWLL
GP*

>ORF21333 (SEQ ID NO:263)
TSGPSRFLVLARRKSSRPTCSRIPNCCSLSSLWTTQASRSSPMPKEVWNDAAAAHRGVHLRARSPGSQREDLPRRDQGAA
QALRSYGNRTGRGPPGCPGMATQGTGTRPVQAELEHVLESSADDLGLCHRA*

>ORF22074c (SEQ ID NO:265)
VRHKPLRYWHYELFVSLLCLAADEDQLIFPVDVALADAQQRIQPDAGEVEDLKRAKPEPGGRDPCAFAGAALCAVLHVEP
IAGEQDGFGGDGFACSPGGNDGGLSERVDLGVRHQLMLDGIAPDRPQMIRVRVPAPLGQALFQYLASPSQDSPVVHVLYG
CRRTEVLEQRLGRGAVDLRAGFPEIVREDVLLGELLRRHSTPPSASVMTVRLGLSRGCSGNSNLGSWSR*

>ORF21421 (SEQ ID NO:267)
AASGQPKPHGHHRCRRRCGMTPQQLTEEYIFAHDLREASAKIYRAATKALLKHFGPTATVQDVDHRAVLGWRRKVLEQGL
SKRSWNTYSNHLRTIWGYAIEHELVTHSQVNPFRKTTVIPPRRASKTVAAEAILLARNWLNMQDGAERCTGERARITPAW
FWLCTFEVFYFTGIRLNALLCIRKRDIDWENQLILIRGETEKTHKEFVVPITEGLVPHLSRLLQEADRAGFADDDQLFNV
NRFSPHYKSKVMNSDQVEAMYRKLTEKVGVRMTPHRFRHTLATDLMKAPERNIHLTKCLLNHSNIQTTMSYIEADYDHMR
AVLHARSLAQGALENVRKVDYSGSPQASAKPKPCGQPLARVSEAPPPEARTEPAEPREHTPGTGIQGGPTAWEADALPQP
PDTFEPSVLFTLMAQNLSNRAASASAAPAATSGSGGWGSAARSNLA*

>ORF22608c (SEQ ID NO:269)
RICFPRGWTSLNACPWRVLPWFCRLCPGLRWRRFTHSSERLPAWLRFGRGLRGAAVIHLPDILQRALGQASSMQHGTHVI
VVGLDVAHRGLDIRVVEQALREVNVPLGCLHQVGGQGVPETVRGHPHPNLLGQLPVHGFDLVGVHHLALVVR*

>ORF22626 (SEQ ID NO:271)
HLRTKRAVHSDGSKLIEPCRLGIRGSRCNKRIRRMGICRPKQSRLAIPVLRAGYRTKGSRAFQQIVRPVGKIWNYRERLD
SSAGMLAEPAQFQGQYHSTLCNR*

Fig. 4-11

>ORF23228 (SEQ ID NO:273)
RDSNSRHPAPKAGALPDCAIPRLEFGSATWTRTRDPMINSHLLYRLSYRGTSFFQPWTLPVLLDSRLRGAPFYGCARACQ
PSDPKSFSSFSTSDKTALPLHAAALSRLPDAHEKAPPKRGFPCPPPKRSGEDDLVAFHLRRDTGTRREFAGQDQLRQRVL
DPALDGPLQRACAIDRVEADGNQLVQRLLAQFQAQLALGQALAQATELDLGDAGDLLASQRLEHHHFVDPVDEFRTEVRI
DRVHHCGTLRLAVAGQLLDLRRTEVGGHHHHGVAEVHRTPVTVGQASVLEHLEENVEYIRMGLLHLVQQHHRVGLAADRL
GQVAAFLEADVARRRADQAGHRVFLHELGHIYPHQRLLGIEEELGQRLAQLGLAHPGRAEEEERAARPVRIGEAGARTAH
GVGHGDYRLVLADHSPMQLLLHAQQLLALALEHLRHRDTGPLGNHFGDFLVGHLVAQQLVLGLAVLVDHLQAAFQVRDGL
VLDARHALEVALAPRRLHLLLGLLDLLLDLRRALHLGLLGLPDLLEVGVFALELDDILLQLGQALPGGFVVFLLQRLALD
LQLDQATVETIQFLRLGVDLHADAAGGLVDQVDGLVRQLPIGDVAVRQLGRGDDRAVGDAHPVVHFIAFLEATEDGDGVF
LARFVHQHLLEAALQRGILLDVLAILVEGSSTDAVQLAARQSRLEHVAGVHGTFRLAGADHGVQFVDEQDDPAFLLAQFV
EDRLQAFLELAAELGTGDQRPHVQGQQALVLEAVRHFAVDDALGQALDDGGLADAGFADQHRVVLGPPLQDLDGPADLVV
ATDHRVELAFLGALGHVDGVLVQRLARLLDVRVVHRFAATQVGHGILQRLARHALAEQQLAEPGVLVHRGQQYQLAGDEL
VALLLGQAVSLVEQACEILGQVHVAGRALDLRQRVEFFVEAAAQGGDIEADLHQQGLDRTALLLEQGGKQVHRLDGRMVM
ANGQGLGVGERQLQLAGQTVYSHGSSFLL*

>ORF23367 (SEQ ID NO:275)
AIAERLSSNPGRFRCCWIRVSEARHFTDARGHVNPLIQKVFLLFPRATKRPFHCMRQRSRAYRTPMKKPRRSGAFPVRPR
RGQAKTISSPSTFAEILAPGANLPARISCASGFSIQRWMARFSGRAP*

>ORF25103c (SEQ ID NO:277)
SAPRGEHHRRRDHRRGQAVAPLHHRSATAGQGHRPDRRGRQPHPHGDRLQAGGTGSSRPSPDPAEDRARGAEEGRRRSHQ
EAPGQAGGGYRQARARIRRPRGDLEVREGRGAGLGADPAEDRAGQAGDGGGAAQGRPREHGAHPVPDHPGPGTQPADGRP
ARQDREPVAAQQGDRRGNRRSGFQVDRYPGVEDARGRAREAAAHGAGAASASDRPGRGGSRRVQRRAPFARRPRRSEPAE
RLVPLPRPDRGGQDRVVQGAGRVPLRYRGGAGADRYVRVHGETLGGPPDRRASGLRRLRGRRLPDRGDPPQALLGGAAGR
GGEGPSGCIQHSPPGARGRTPDRQSRAYGGLPQHRGGDDLQPRFGADPGAGRRPRGATCRSDGRGQCALPSGIHQPDRRS
GGVRAAGSRADRRHRRDPARSPAQAPGRARAEPGTEPGGAGQADCRRLRPGLWRTPAEAGHPALDREPAGATDPGRQIRA
GCQYLGEGGRRRDRLRLTSSGADRESPASAGLFHGRPVGARALPHAVEGPFCRSWKKKKNFLDQRVDMPARIRKMARL*

>ORF23556 (SEQ ID NO:279)
KSPAEAGLSLSAPEEVRRRRSRRLPPSPRYWHPARICRPGSVAPAGSRSSAGWPASAGVRHRPGRSRRQSACPAPPGSVP
GSARARPGACAGDRAGSRRCRRSAREPAARTPPLRRSG*

>ORF26191c (SEQ ID NO:281)
KEGRPMRIDRLTSKLQLALSDAQSLAVGHDHPAIEPVHLLSALLEQQGGSIKPLLMQVGFDIAALRSGLNKELDALPKIQ
SPTGDVNLSQDLARLLNQADRLAQQKGDQFISSELVLLAAMDENTRLGKLLLGQGVSRKALENAVANLRGGEAVNDPNVE
ESRQALDKYTVDMTKRAEEGKLDPVIGRDDEIRRTIQVLQRRTKNNPVLIGEPGVGKTAIVEGLAQRIINGEVPDGLKDK
RLLALDMGALIAGAKFRGEFEERLKAVLNELGKQEGRVILFIDELHTMVGAGKAEGAMDAGNMLKPALARGELHCVGATT
LDEYRQYIEKDAALERRFQKVLVDEPSEEDTIAILRGLKERYEVHHGVSITDGAIIAAAKLSHRYITDRQLPDKAIDLID
EAASRIRMEIDSKPEELDRLDRRLIQLKIEREALKKEDDEATRKRLAKLEEDIVKLEREYADLEEIWKSEKAEVQGSAQI
QQKIEQAKQEMEAARRKGDLESMARIQYQTIPDLERSLQMVDQHGKTENQLLRNKVTDEEIAEVVSKWTGIPVSKMLEGE
REKLLRMEQELHRRVIGQDEAVVAVSNAVRRSRAGLADPNRPSGSFLFLGPTGVGKTELCKALAEFLFDTEEALVRIDMS
EFMEKHSVARLIGAPPGYVGFEEGGYLTEAIRRKPYSVVLLDEVEKAHPDVFNILLQVLEDGRLTDSHGRTVDFRNTVVV
MTSNLGSAQIQELAGDREAQRAAVMDAVNAHFRPEFINRIDEVVVFEPLAREQIAGIAEIQLGRLRKRLAERELSLELSQ
EALDKLIAVGFDPVYGARPLKRAIQRWIENPLAQLILAGKFAPGASISAKVEGDEIVFA*

Fig. 4-12

>ORF23751 (SEQ ID NO:283)
TGSKPTAISLSSASWLSSRLSSRSARRLRRRPSWISAMPAICSRASGSNTTTSSIRLMNSGRKCALTASITAARCASRSP
ASSWICAEPRLEVITTTVLRKSTVRP*

>ORF24222 (SEQ ID NO:285)
PGGAPIRRATECFSMNSDISIRTSASSVSKRNSASALHNSVLPTPVGPRKRNEPLGRFGSARPARERRTALDTATTASSW
PITRRCSSCSMRSSFSRSPSSIFDTGIPVHLETTSAISSSVTLLRSNWFSVLPCWSTICRLRSRSGMVWYWMRAMLSRSP
LRRAASISCLACSIFCWICAEPCTSAFSDFQISSRSAYSRSSLTISSSSLARRFLVASSSSFFSASRSIFSWIRRRSRRS
SSSGLESISMRMRLAASSIRSMALSGSCRSVM*

>ORF24368 (SEQ ID NO:287)
TRTYLSAPAPPRYRRGTRPAPCTTRSCPPRSGRGRGTSRSAGSDRRGRRANGARRWTRRLPPRPGRSLADAAPAPCAAAS
RARPRASSTPGYRSTWKPLRRFPRRSPCCAATGSRSCRAGRPSAGCVPGPGWSGTGCAPCSRGRPCAAPPPSPAWPARSS
AGSAPSPAPRPSRTSRSPRGRRIRARA*

>ORF24888c (SEQ ID NO:289)
RRKTTKPPGSAWPSWRRISSSSSANTPTSRRSGSPRRPRCRARRRSSRRSSRPSRRWRRRGARATSRAWRASSTRPSRTW
NAACRWSTSTARPRTSCCATR*

>ORF25398c (SEQ ID NO:291)
RRSSTNWASRKAGSSCSSTNCTPWSAPARRKVPWTPATCSSRLWRAASCTASVLLPSTSIASTSRRMPRWSAASRRCWWT
NRARKTPSPSSVASRNAMKCTTG*

>ORF25892c (SEQ ID NO:293)
PPGPAEGRPVHLQRAGIAGRDGREHQARQAAARPGRVAQGAGECRGQPAWRRSGERPERRGVAPGAGQVHRRHDQARRGR
QARPGDRSRRRDPPDHPGPAAADQEQPGADRRTRRRQDRHRRGPGPAHHQRRSAGRPQGQAPAGPGHGGADRRCQVPRRV
RGTPEGGPQRTGQAGRPGHPVHRRTAHHGRRRQGGRCHGRRQHAQAGSGARRAALRRCYYPRRVSPVHREGCRAGAPLPE
GAGGRTERGRHHRHPPWPQGTL*

>ORF25110 (SEQ ID NO:295)
RSLRPRRMAMVSSSLGSSTSTFWKRRSSAASFSMYWRYSSRVVAPTQCSSPRARAGLSMLPASMAPSALPAPTMVCSSSM
NRMTRPSCLPSSLRTAFRRSSNSPRNLAPAISAPMSRASRRLSLRPSGTSPLMMRWARPSTMAVLPTPGSPISTGLFLVR
RCRTWMVRRISSSRPITGSSLPSSARLVMSTVYLSSAWRDSSTFGSFTASPPRRLATAFSSALRDTPWPSSSLPSLVFSS
IAASNTSSLEMNWSPFCWARRSAWLSRRARSWDRFTSPVGLWIFGSASSSLLRPLRRAAISKPTCISRGLIEPPCCSSRA
ESRCTGSMAGWSWPTARDWASESASCSLLVKRSIRMGRPSFYRAGRNDGCP*

>ORF25510 (SEQ ID NO:297)
CAGPGPRRWRSCRRRVRRSAPGCSWSAAAGPGWSGGSRRRDRSPGRACLPRRAWSCRRCTCPAPGATPRRSGRSPLRRHA
GWPRHSPAPCATRPGRAAACRAWCSRPSRPAIPARWR*

>ORF26762c (SEQ ID NO:299)
PPTACRRCSATARAPGWPRPMPAGAGWRRACWRRRWTAWACPATNCWSGWGRRSARRPSRSAARSAMHSSLRTPRRARLS
YLAPIRAASWPTSTDSRGSAWAPMASPPCMAAASAPSAIPRASIPTAARRVPAVLPAWSGSRTRPAQVIRRQLTDVTVRS
LEPRKIALIY*

>ORF26257 (SEQ ID NO:301)
IRAIFRGSSDRTVTSVSCRRITCAGLVLEPDQAGKTAGTRRAAVGIEARGIAEGAEAAAMHGGDAMGAQADPRESVDVGH
EEAARIGARYESRARLGVRSDECIADLAADLEGLRADRRPQPDQQFVAGHAQAVHRRLQHARRQPAPAGMGRGHPGARAVA
EQRRQAVGGHDRTGDARHRAPAGVGPEHRFGSASTTSLRAPIPTSSAGIPDARSGVGGFIPRRADRRRRGRQGSDCRRVP
G*

Fig. 4-13

>ORF26844c (SEQ ID NO:303)
RGGGRPEPVLRADASWSAMPGVACTIMTADCLPALFCDRSGTRVAAAHAGWRGLAAGVLEATVDSLGVPGDELLVWLGPA
IGPQAFEVGGEVRDAFVAAHAEARSAFVPSANPGRFMADIYRLARIRLGAHGVTAVHGGGFCTFSDTARFYSYRRSSRTG
RFASLVWLQD*

>ORF26486 (SEQ ID NO:305)
MSAMKRPGLALGTKAERASACAATNASRTSPPTSKACGPIAGPSQTSSSSPGTPRLSTVASSTPAASPRQPAWAAATRVP
ERSQNNAGRQSAVMIVQATPGIALQLASALSTGSGRPPPRHSVHLFQPARPAFQTLGQASAVLFHGARIVVDVGAKVQTV
EGCLADPATARGHAGPHTGRRRPVGGQPGVQPTNASRSWRNRLSSQRNSSGSGDSHFMRTPVAG*

>ORF26857c (SEQ ID NO:307)
VHGVTWWRPTRTGAQGRRQLERDAGRRLYDHDRRLPAGVVLRPLGHPGGRGPCRLARAGGGRAGGDGGQPGRARRRTAGL
AGAGDRPAGLRGRRRGPRCIRRCARRGALGFRT*

>ORF27314c (SEQ ID NO:309)
SGNRRCRKNSSGCSACCARIARRSSVERLADPRLAGAGPCAGLRDHAQWRGQPGTLRQSEPWRPRLRRSARRGIKPPTPD
RASGMPAELVGIGARSDVVEADPNRCSGPTPAGARCRASPVRS*

>ORF27730c (SEQ ID NO:311)
QARRPGGPSGCRPSGRHPAECLALPCPGHRQCAARRDRPPPGQGHDRPDGSGQDAGGPHQAGGATAGTVGQPHLRGDRDR
RDHLRRHHRCADRTAWRAAAEDGGGRRRQGGGQPLPRAGTLPCAHPYPGQAGDRAYPPDPRAHEPYWLSPGRRSGLRWAL
QDSPGGQPDPGPDSSRIPPAGAAARALPRTGSPGHRRAHEVGIAAAGRIPLAAQPVAPGSRGVRRLNAWLTPDWPAPARVR
ACVTTRSGGVSQAPFDSLNLGAHVYDDPRAVE*

>ORF26983 (SEQ ID NO:313)
PRHCAWSRRPAHGPAPASRGSARRSTDERLAILAQQAEQPEEFFRQRRFPLHAHAGGRVIQFEEARVQRLPGEFAKSLDQ
GLAGHRGNPEAPTVDRIADQGIANMAHVHADLVGTPGLQLDPGMGVRTEAFQHAVMADRHLAGVDHRHLLPLHAMPSDRR
IDGAAGGDHADHDRLVDAADRPCLQLRHQLGVGLQRLGHYHQAGRVLVQAVDDPGARHIGDVRDMVEQGIQQGAVLMAGS
RMDHQAGGLVNHQDVLVLVDDFQLDVLCEPLALGFLLGLQDQLRAAVDDVARAQHGAVDGQATVLDPAGQTGAGVFGKKL
GGDLVETLATQLERHLGRALNHIGHE*

>ORF28068c (SEQ ID NO:315)
PQRVADSKSRAEHRLLLMSDMIQRAAEVPFELGGQRLDQIAAQLFPEHSRSRLAGWIKDGRLTVDGAVLRPRDIVHSGAQ
LVLEAEQEAQGEWLAQDIELEIVYEDEHILVIDKPAGLVVHPAAGHQDGTLLNALLYHVPDIANVPRAGIVHRLDKDTTG
LMVVAKTLEAHTKLVAQLQARSVSRIYEAIVIGVITSGGTIDAPIGRHGVQRQKMAVVDAGKVAVSHYRVLERFRAHTHT
RVKLETGRTHQIRVHMSHIGYPLVGDPVYGGRFRIPPVASQTLVQTLREFPRQALHARFLELDHPATGVRMKWESPLPEE
FLWLLSLLRQDREAFVG*

>ORF27522 (SEQ ID NO:317)
PTVPAVAPPAWCGPPASWPLPSGRSCPCPGGGRSRRAAHWRCPGHGRARHSAGCRPDGRQPDGPPGRRACQSPGCARPRR
RFPARCPVRATRPGLPARPPGPVARRCGRCRAGAARRRRRSGDRP*

Fig. 4-14

>ORF28033c (SEQ ID NO:319)
ASSPTHVRYDSTRGRGAVRAGWPASRPDRRPAFSRTLPLPSGRLDQGRSPDRRRRRAAPARHRPQRRATGPGGRAGSPGR
VARTGHRAGNRLRGRAHPGD*

>ORF29701c (SEQ ID NO:321)
SSSSLEISRTSTRPMVRRYRWYRRRMRCPCSSLSRSRSARTVALVLAQVRLAAIPALFVGEGVGLRHVDAAMGAADHRRC
ARLVLRTLLLARSGTGKATPEPERDGDQGDPEQEAEKAHGDLGGWRKLQFSQAAGSIPDGKVQAVRRLALGEAWRRAKRR
EACASLRCFVRSVEETEGHVAPPGATGVLVVIALRLLVVGAVILVFRLQFGGDLPLGILVLLDHVLGGLGFHVRRRLAAF
DQAQGGLGQPGAGVGLAFAGDELAILEAGVIRIVQLEGFQAGAGQVVETQATVGFDHDRQAIADGRGFLEVLHHVATAVG
GGDIGLALQVVVADVHFVGRQQVAQVHHARLGVRGVAAVGEAAGELGELVEGVAGGARVALGHVQRQEARQQAAVLVEGG
QAFEVVGVVDVGVLRMQADEAFGGGAGGFGLHVLVVVGVDQLELGLLGVAAEGIARFEGFQLGDGAVVALVVEVVLRLLVQ
LALAQVLVDSLLVRGAGCGEGEDGDQQQVFHLHGGLRPWDGRLGLNRLL*

>ORF28118 (SEQ ID NO:323)
QTVEAKPTVPGTQAAMQVKHLLLIAILALTAACSSNKETVDENLSESQLYQQAQDDLNNKSYNSAVTKLKALESRYPFGR
YAEQAQLELIYANYKNMEPEAARAAAERFIRLHPQHPNVDYAYYLKGLSSFDQDRGLLARFLPLDMTKRDPGAARDSFNE
FAQLTSRFPNSRYAPDAKARMVYLRNLLAAYEVHVGHYYLKRQAYVAAANRGRYVVENFQETPAVGDGLAIMVEAYRRLG
LDDLASTSLETLKLNYPDNASLKDGEFVARESEADTRSWLAKATLGLIEGGEPPPHMETQAAKDVIKQYEDAEREIPAEL
KPENQDHSADDEKPESDDDEDSGRSWWSYMTFGLFD*

>ORF28129 (SEQ ID NO:325)
GQADRPRDASRHASETPAADRHPRPHRSLLLEQGDCRREPEREPAVPAGAGRPQQQELQQRRHQAESPRIALSLRPLRRA
GPARADLRQLQEHGARSRPRRRRTLHPPASAAPQRRLRLLPQRPVLLRPGPRPAGALPAAGHDQARPGRRPRLLQRVRPA
HQPLPQQPLRPGRQGAHGVPAQPAGGLRSARRPLLPEAPGLCRRRQPRSLRGGELPGNPGRRRWPGDHGRSLPSPGSRRP
GQHQPGNPQAELSG*

>ORF29709c (SEQ ID NO:327)
GPDLPVRWRSAGPVPGRWSGGTGGTDGGCVAPAPRSAVAVQRARSPWSWRRCAWLQYQRCSWARALACGTWTPQWAQRTI
GDALGSSCGRCCWRGVGRVKRRQSQNAMAIRAIQNRRRKRPMVISEAGESCSLAKPPARSQTGRSRLCGVWRWERHGGGQ
KEGRPAPPFGVSCDQSKRPKVM*

>ORF29189 (SEQ ID NO:329)
SHETPKGGAGLPSFCPPPCLSQRQTPHSLDLPVWDRAGGLAKLQLSPASEITMGLFRLLFWIALIAIAFWLWRRFTRPTP
RQQQRPQDEPSASPMVRCAHCGVHVPQANALAHEQRWYCSQAHLRQDQGDRAR*

>ORF29382 (SEQ ID NO:331)
SPSRSGSGVALPVPLRASSNVRRTSRAHRRWSAAPIAASTCRRPTPSPTNNAGIAARRTCARTRATVRAERLRLSEEQGQ
RILRLYHLYRLTIGLVLVLLISSELEDQVLKLVHPELFHVGSWCYLVFNILVALFLPPSRQLLPIFILALTDVLMLCGLF
YAGGGVPSGIGSLLVVAVAIANILLRGRIGLVIAAAASLGLLYLTFFLSLSSPDATNHYVQAGGLGTLCFAAALVIQALV
RRQEQTETLAEERAETVANLEELNALILQRMRTGILVVDSRQAILLANQAALGLLRQDDVQGASLGRHSPMLMHCMKQWR
LNPSLRPPTLKVVPDGPTVQPSFISLNREDDQHVLIFLEDISQIAQQAQQMKLAGLGRLTAGIAHEIRNPLGAISHAAQL
LQESEELDAPDRRLTQIIQDQSKRMNLVIENVLQLSRRRQAEPQQLDLKEWLQRFVDEYPGRLRNDSQLHLQLGAGDIQT
RMDPHQLNQVLSNLVQNGLRYSAQAHGRGQVWLSLARDPESDLPVLEVIDDGPGVPADKLNNLFEPFFTTESKGTGLGLY
LSRELCESNQARIDYRNREEGGGCFRITFAHPRKLS*

>ORF30590c (SEQ ID NO:333)
LLQQLGGVADRAQRVADLMGDAGGQAAKTGQLHLLRLLGDLRNVFEEDQHVLVVFAVEADKAGLHRRAIRHHLERRRTEA
GIQAPLLHAVHQHRAVAAEAGALHVVLPEQAEGGLVGEEDGLTAIDHEDAGAHALQDQCVEFLQVGDRLGAFFGQRFGLL
LAPHQSLDHQRGGEAQGAEEAAGLDVVVGGVRTAQAEEEGQVEQAEAGRRRDDQADAPAQQDVGNGHRHHQQAADAAGYAA
TCVEQAAKHQHVGEREDEDRQQLPRRRQEQRDQDVEDQVAPTADMEQFRVDELEDLIFQFAGDQQDQYQADGQAVQVVQT
EDALPLLLAQP*

Fig. 4-15

>ORF29729 (SEQ ID NO:335)
TVPCRQLVLPGLQHPGRAVPAAVAAIAADLHPRAHRRADALRPVLRRWRRTQRHRQPAGGGGGHCQHPAARAHRPGHRGG
GQPRPALPDLLPQPEQSGRHQPLRPGRRPRHPVLRRRAGDPGSGAAPGADRNAGRRTRRDGRQPGGTQRIDPAAHAHRHP
RGR*

>ORF30221 (SEQ ID NO:337)
PSGHPPRQPGRPRPAQAGRRAGRQPRPPQPDADALHEAMAPESQPPSADAQGGAGWPDGATQLYQPQPRRRPARADLPRR
HFADRPAGAADEAGRSWPPDRRHRP*

>ORF30736c (SEQ ID NO:339)
SHSFRSSCCGSAWRRRESCRTFSMTRFIRFDWSWMICVRRRSGASSSSDSCSSWAAWLIAPSGLRISWAMPAVRRPRPAS
FICCACWAICEMSSRKISTCWSSSRLRLIKLGCTVGPSGTTLSVGGRRLGFRRHCFMQCISIGLWRPRLAPCTSSCLSRP
RAAWLARRMA*

>ORF30539 (SEQ ID NO:341)
DPQPAGRDQPRRPTAAGVRGTGCPGPTPDADHPGPVEADEPGHRERPAALPSPPGRTAAARPEGVASAVRRRIPRQAAQR
QPTAPAARCRRHPDPHGPTPVEPGAEQPGAERSSLQRPGARARPGLAEPRARPGERPAGAGSHRRRSRRTGGQTEQPVRT
LLYYRKQRHRPGPLSLPRTLREQPGTDRLPQSRGRRRLLPHHLRPPAQTQLTEAARMSRQKALIVDDEPDIRELLEITLG
RMKLDTRSARNVKEAASCWPASRSTCASPTCACRTAAASIWSSTSSSAIHRPRWP*

>ORF31247c (SEQ ID NO:343)
FPAVRGYPVHRRRSGLFVGSCVRLPSAEFARVGEGDAEAAAAFLAIAVVDPCLVALAEFAGEIEAQAGAFAFCSKEGFEQ
VVQFVRRYAGTVVDDFQHRQVALRVAREAQPDLAAPVRLGAVAKTVLHQVAQHLVQLVWVHAGLDVAGTELQVQLAVVAQ
PAGVFVDEPLKPLLQVELLRFGLAATGELQDVLDDQVHPLRLVLDDLRQASVRGIQFL*

>ORF30963c (SEQ ID NO:345)
LPAPAGRSPGRARGSARPGRARAPGRCSEDRSAPGCSAPGSTGVGPCGSGCRRHRAAGAVGCRCAACRGIRRRTAEATPS
GRAAAVRPGGDGRAAGRSR*

>ORF31539c (SEQ ID NO:347)
GGCHQLPQATEVDRFGEEVEGTGLERLDRGVQAAVRGDHGHRGLWMALLDVLDQIEAAAVRQAHVGEAQVERLAGQQLAA
SLTLRALRVSSFMRPRVISSSSRISGSSSTIRAFCRLMRAASVS*

>ORF31222 (SEQ ID NO:349)
TGYPRTAGNHSRPHEAGHPQRPQRQGSRELLAREPFDLCLTDMRLPDGSGLDLVQYIQQRHPQTPVAMITAYGSLDTAIQ
ALKAGAFDFLTKPVDLGRLRELVATALRLRNPEAEEAPVDNRLLGESPPMRALRNQIGKLARSQAPVYISGESGSGKELV
ARLIHEQGPRIERPFVPVNCGAIPSELMESEFFGHKKGSFTGAIEDKQGLFQAASGGTLFLDEVADLPMAMQVKLLRAIQ
EKAVRAVGGQQEVAVARAHPLRHPQGPRRRSRRRALPPGPLLPPQRHRAARTPLRERREDIPLLAERILKRLAGDTGLPA
ARLTGDAQEKLKNYRFPGNVRELENMLERAYTLCEDDQIQPHDLRLADAPGASQEGAASLSEIDNLEDYLEDIERKLIMQ
ALEETRWNRTAAAQRLGLTFRSMRYRLKKLGID*

>ORF31266 (SEQ ID NO:351)
SWTPAAPATSRKPRVAGPRAVRPVPHRHAPAGRQRPRSGPVHPAAPSTDPGGHDHRVRQPGHRDPGAQGRCLRLPHQTGR
PRSLAGAGGNRPTLAQPGSRGSAGGQPPARRVAADARPAQPDRQAGAQPGAGLHQWRVRQRQGTGGAPDPRAGATYRAAV
RAGELRRDSLRADGKRVLRPQERQLHWRYRRQAGPVPGRQRWHPVPRRSRRPADGHAGQTAPGDPGKGRARGRRPAGGRR
RTCASSAPPTRTSPPKSAPGASARTSTTASTSSSCAYTAARTPRGHPAARRTHPQAPGRRHRPAGRQADRRRTGEAEELP
LPGQRPRAGKHAGARLYPVRRRPDPASRPAPGRCAGCQPGRRREPERNRQPRGLPGRHRAQADHAGTRGDPLEPHRRGPA
PGPDVPLDALPPEKAGHRLKVKRPVRRQAFWFSLLRGDQPGRRGPGR*

Fig. 4-16

>ORF31661c (SEQ ID NO:353)
TGAWLRASLPIWLRRARIGGDSPSRRLSTGASSASGLRKRRAVATSSRKRPRSTGLVRKSKAPALSAWIAVSRLPYAVIM
ATGVCGWRCWMYWTRSRPLPSGRRMSVRHRSNGSRASNSRLP*

>ORF32061c (SEQ ID NO:355)
RSWRKRPAPTSAARSLWVAQRMRTCDGDLLLAADRAHGLFLDRPEQFDLHGHRQVGDFVEEQGATAGGLEQALLVFDSAS
EAAFLVAEELAFHQLGGNRAAVHRHERPLDTWPLLVDQARHQFLAAAGLATDVDRRLAARQLADLVAQGAHRRRLAEQAV
VHRRFLGFRVAQA*

>ORF32072c (SEQ ID NO:357)
GGSRGPGGSARRRLRRRGPCGWRRGCARATATSCWPPTARTAFSWIARSSLTCMAIGRSATSSRNRVPPLAAWNRPCLSS
IAPVKLPFLWPKNSLSISSEGIAPQFTGTNGRSIRGPCSWIRRATSSLPLPDSPLM*

>ORF31784 (SEQ ID NO:359)
WKASSSATRKAASLALSKTSRACSRPPAVAPCSSTKSPTCRWPCRSNCSGRSRKRPCARSAASRRSPSHVRILCATHKDL
AAEVGAGRFRQDLYYRLNVIELRVHRCANAARTSRCSPNASSSAWPATPACRPPG*

>ORF32568c (SEQ ID NO:361)
GAKTKRPVFGQAFSLSVDAQLFQAVAHRAERQAQALGRGGAVPAGLLECLHDQLALDVFQVVLEVVDFAQARGAFLAGTR
RIGQAQVVRLDLVVFAQGIGALQHVFQLADVAREAVVLQLLLCVAGQPGGRQAGVAGQALEDAFGEQRDVLAAFAQRCTR
SSMTLRR*

>ORF33157c (SEQ ID NO:363)
TDGGARLVARRRSSGRRGDLAWRDPWRQGAAGGRRLERRVVEAAWPGTARGTGERSDDPLQVRGGFPAAHGAGQGALRDS
AARRPHPDRQHLGTFGLRQDADRRGAGKPQGVCGRTVAGTGGHAAGGPLGRVAPGLSRRHPLYRSGAWLRRALAEYRALP
QRAGPGTGVVPSAGGSHERAGTDHRPGPLRPGWSPLRSENQKACLRTGLFTFSRCPAFSGGSASSGTSGPGAGPRRCGSS
GSPRVPA*

>ORF32530 (SEQ ID NO:365)
KGLSEDRPFGFRSSEATSRGVGGRVDDRFPPAHEIRQQTARRRCQDQPVAVVPGIQPEPVEARHRTDIGDAFGRARAQPC
PVGHRLHVRQFRQQFCRRRPEAFQRLVGRRLVEARMFQGAADQDVAVAPRNRVAPLGQHHARQEIRRALVEDHLTFHRYH
GQFQAKRLQQLAAPGACRQQHLVATDLATRGRHADHSIAVAQPAAHLRLFMQLEIGELLQGCPQ*

>ORF33705c (SEQ ID NO:367)
VIFLCSWQIGRSPVVSRDVVVVGAGVIGLLTARELALAGLRVTLVERGESGREASWAGGGIVSPLYPWRYSPAVTALAHW
SQDFYPALGQRLLDETGLDPEVHTVGLYWLDLDDQTEALQWARNHTRPLKEVPIEEAYAAVPGLGAGFQRAVYMSGVANV
RNPRLARSLRASLQQFANLELHEQTEVRGWLRDGDRVVGVATSRGEIRGDKVLLAAGAWSGELLKPLGLELPVVPVKGQM
ILYKCAADFLPRMVLAKGRYAIPRRDGHILIGSTLEHSGFDKTPTDEALESLRASAAELLPELADMQPVAHWAGLRPGSP
EGIPYIGPVPGFDGLWLNTGHYRNGLVLAPASCRLLADLMSGREPIIDPAPYAPAGRL*

>ORF32832 (SEQ ID NO:369)
GFPAPRRSASCRSPNVPRCCRSGCGRRAAESRSAPWPAPCAAGNPPRTCRGSSDLSPVPRAVPGQAASTTRRSRRLPPAA
PCRHGSRHARSPRRPLDRRRATSRAPPSVHATRDWRIVAGMPAMSAPGEDCAHWPRPTCRPPAGSLRPARAPPRRPPRSA
LPSTAGCGCVPTAVPRSGRPGPASTGQRYGPRDRARSRRANAAPGPGRSPATSAPGRSPPGCSATDRAARRSRLPPRMPH
AHSRPAPPGSPAVRRAPAPGRSTGR*

Fig. 4-17

>ORF33547c (SEQ ID NO:371)
GILGGRRDRLAALSVALQPGGDRPGALVAGLLPGPGAAFARRDRARSRGPYRWPVLAGPGRPDRGTAVGTQPHPAVEGSA
DRGGLRGGARAGRRLPAGGLHVGRGQCAQSSPGALIAGIPATIRQSRVA*

>ORF33205 (SEQ ID NO:373)
ARQARIAHIGHARHVDRPLEACAQPGHRRVGLLDRHFLQRPGVVACPLQCLGLVVQVQPVQANGMDLGIEPGLVEQTLPQ
GRVEVLRPVRQGGHRRAVAPRIERRDDPASRPGCLTPTLAPLHQGHPQSGERQLPGGQQADDASAYHYYISTHHRAPTDL
PGTEKYHSKGSDADELPASTNSVESSPGEKPIIPAEVFIP*

>ORF33512 (SEQ ID NO:375)
SGETIPPPAQDASRPLSPRSTRVTRSPASASSRAVNRPMTPAPTTTTSLLTTGLLPICQEQRNITQRDQMLTNCLLQRTQ
SNLVPVKSPSYPQRYSSHEIEWFEFGGTIDSPIDPCDRRDNCAAHPPRQNEAGH*

>ORF33771 (SEQ ID NO:377)
KAHHTRRGIHPMKSSGLNLVELSIVLSILAIGVTIALPTLPDRMKRDISRDIGDSLTSHVMAARASSIQNGVIIEVCGSG
DGSTCSEEWHLGWFSRNDRSQQILARHENTSRTDIHWRGFDKRLRYLPNGTSPTGNGRFFECKDDRIEWQLVLNRQGRLR
VAGKSENKKLSYLCSRR*

>ORF34385c (SEQ ID NO:379)
WRAHCPCSRDSRSGEWDDWQTVCETVLSPPGAQVRELFIFALSRHPEAALPIEHQLPLDAIVLTFEETPVTCRAGAIRQV
AQSLVEAPPMNIGATRIFMPGQYLLAPVVTAEPAEMPFLAAGTAVTATAHLDDHAVLYAASPRSHHMTSQAVTNITANVP
LHSVGEGGQRNCHAYRKDR*

>ORF33988 (SEQ ID NO:381)
SSRCAVAVTAVPAARNGISAGSAVTTGANRYWPGMKIRVAPIFIGGASTSDCATCLMAPALQVTGVSSNVRTIASSGNWC
SIGKAASGWRERAKIKSSLTCAPGGERTVSHTVCQSSHSPLRLSLLQGQCALH*

>ORF34274 (SEQ ID NO:383)
KALLPVLQAVRELFHIPFASHPTLRSGCLCYRDNALSTRQDYLALFLVEYCMRSICRSAGFSLIELMMVLVLVAIFASIA
VPSFNALIERNRIQTASEELYSLLQYARSEAVNRHANVSIRATQNNDWAKGLEIISGATTVQKHQGFQQVSLSASSATAE
LTFNATGTLSNQAANIDIKVCFAGDKSTGRLLTVQPSGRVILYPSSKQPDSCN*

>ORF34726c (SEQ ID NO:385)
RDLLETLVLLHGGRAADDFQAFCPVIVLRRPDAHIGMTVYSFAASVLKQAVEFLAGSLDSVALDQGVETGYGNAGEYGDQ
NQHHQLDQGKAGAATNRAHAVLHKEKGQIILPSGERIVPVAETAGAESGMTGKRYVKQFSHRLEHR*

>ORF34916 (SEQ ID NO:387)
GKPMSRETGFSMIEVLVALVLISIGVLGMVAMQGRTIQYTQESVQRNAAAMLASDLMEIMRADPDAVLNLRAQLREDSVY
YKAKGSDFPAAPARCAPLPADAKERLGCWAQQASKDLPGASALLNSQFYICRSPTPGTCDNTKGSAIEIQVAWRAMDGAC
FNASDSTLCTYSVRSEL*

>ORF35464c (SEQ ID NO:389)
RACLFSQFGADAVGAQGGVRGVETRSIHGSPGNLDFDGRAFGVVAGTRGWAATNVELAIQECGGSRQVFRGLLGPAAETF
LSICWQWRAARWGCGKVAALGLVVDRVFA*

>ORF35289 (SEQ ID NO:391)
IANSTFVAAQPRVPATTPKARPSKSRLPGEPWMERVSTPLTPPCAPTASAPNCENKHALQQNAERPIDGRTARGTRYKQL
PDPGDQPDLHRQQTQLSFPARPGRQPGK*

Fig. 4-18

>ORF35410 (SEQ ID NO:393)
LHLVHLQRPLRIVRTSMLFSKMQKGLSMVELLVALAISSFLILGISQIYIDNKRNYLFQQGQAGNQENSRFVLMLLQQQL
DKTAYRRLHDDNMENAFKSATFNGCRAFVAGETIAAATALKAGEYGVCLRYQPAYKGEHDCLGNEITGVPEKPFTNTPPV
VVRLVYLPSAGTLSCSRPDIAQSKSGELVSGLTDFRLEAGVGPADRSERKVSSFVALQDVAGRPIRALRFSILAGSDNTS
LRTGDDSQARDRWIVLYPESKSAIEAADKGQIYQIARGNQTIRNLMP*

>ORF35907c (SEQ ID NO:395)
VDQANDDRGSICEGLFRNSGNFITEAIMLPFVGGLIAQADTVLTRLEGSCRSDSLASHKCTTAIECRGFESILHVVVVKA
TIGCLIQLLLQQHKNEAAIFLVAGLALLEKIVAFVVDVDLADPQDQEAAYSECHEQFYHR*

>ORF35534 (SEQ ID NO:397)
SWGSARSTSTTNATIFSSKARPATRKIAASFLCCCSNNWIRQPIVAFTTTTWRMLSNPRHSMAVVHLWLARLSLRQLPSR
RVSTVSACAINPPTKGSMIASVMKLPEFRKSPSQILPLSSFAWSTYRAPVP*

>ORF35930 (SEQ ID NO:399)
VAVVPISPSRNRENWSVVSQTSAWKRGSGQQIVANAKYPASSHYRMSPVVLSEHCASQSWQAATIQACAQEMIARHAIAG
SSFIPRAKAPSRPQTKARFTK*

>ORF36246 (SEQ ID NO:401)
PNHQESHAMTLRHTSRQQGSTLLISLVILLMITLLAVSNMREVSLESRITGNLIEQKRLRNAGEAGLREGERRFFNTIKP
PEVGSGCADSNVKRPCILNLSALSVPRDDVHNNPVAALNGKTDNANSRVWMPYRGSDLNNPTQIDKDRAVTWQTITVPAG
EQNNEAENPEYGNMMRGVGTFYYETNSRALNKAGGETVLQAVHARLYTN*

>ORF26640c (SEQ ID NO:403)
GIQTRELALSVLPFRAATGLLCTSSRGTERALRFSMQGRLTLLSAHPLPTSGGLMVLKKRRSPSRSPASPALRRRFCSMR
LPVIRLSSDTSRMLETARSVIINKITSEINNVDPCCREVWRRVMA*

>ORF36769 (SEQ ID NO:405)
CAGSARSTTKPTAAPSTRRAERLFYRPFMHACIPTDWSQRMIHQITRAGKSLLAAGCTLSILFASDSYAATALNVSQQPL
FLTQGVAPNLLFTLDDSGSMAWAYVPDGISGNSGRAGRSSDYNALYYNPDYAYQVPKKLTLSGDQIIVSDYPVPRFTAAW
QDGYAQGSTTNLSNNYRPQWGTGWLGCIDSSCNTGRAYYYTYKVSASCPAQPVSSSNSCYTYNALPTSQESNFAIWYSYY
RNRILATKTAANLAFYSLPENVRLTWGALNTCSIGANSRSCQNNALLQFNKQHKINFFNWLANSPASGGTPLHAALDRAG
RFLQTNGTAYTTEDGKTYSCRASYHIMMTDGIWNGRNVTPGNLDNQNQTFPDSTLYRPQPPYADSNASSLADLAFKYWTT
DLRPSIDNDLKPFMAYKSGDDSKDYWDPRNNPATWQHMVNFTVGLGLSYSLTLNSAPTWTGSTFGNYEELMAGSKAWPSV
DNDAAPGNVYDLWHAAINSRGDFFSAESPDSLVQAFNKILTRISERNTSSSKPAMTSALQDDGTGDKLIRYSYQSSFASD
KNWAGDLIRYKVESTSTGSTKTQEWSAGALLDNRAPATRNIYIASNSGTNRLKPFTWSNIEGSQLATWLNRNPDKDNQAD
TKGAQRVDFIRGQQNMDGFRQRQAVLGDIVHSSPAVVGPAQYLTYLANPIEPSGDYGTFKTEADQRSPRVYVGSNDGMLH
GFNIKTGVEEFAFIPTAVFEKLNKLTGISYQGGAHQYFVDATPVVSDAFFDGAWHTVLIGTLGAGGRGLFALDVTKPDDV
KLLWEYDSSTDSDLGYTFSKPTVARLHSGQWAVVTGNGYGSDNDKAALLLIDLKKGTLIKKLEVQSERGIANGLSTPRLA
DNNSDGIADYAYAGDLQGNIWRFDLIGNTRNDDPDTNTSINPFKPGDVDPSAFRVSFSGAPLFRARADNNTRQPITAPPT
LVRHPSRKGYIVIVGTGKYFEDDDAQADTSRAMTLYGIWDRQTKGESANSTPTIDRNALTAQTMTTEANSTFGSVNRNIR
LISQNPVKWYKDGATGTANSDVASYGWRLNLEVNSSKKGEMMIEDMFAAGQVLLLQTLTPNDDPCDSGSTSWTYGLNPYT
GGRTSFTVFDLKRAGIVDSGSDYNGSVVSAFQQDGLGGLAITQNEQRQSEACTGDECIIFNPSDKSNGRQTWRVVEEK*

>ORF37932c (SEQ ID NO:407)
AGIAVGIRGLWPIEGAIRKGLVLVVEIAGGDVPTVPDTVGHHDVIAGPAGICLSVFGGISCAVGLQEASGSVKSRMQRST
AAGRAVRQPIEEIDFVLLVELEQGIVLAASAVGADATGVQGPPSETHVFRQAVKGQVGSGLCGQDAVAIVGVPYRKVAFL
TGRKSIVGITGVGAAHRLCRAASAYLISIVISSPGIAAAIDATKPAGSPLRAIVIAQVGGGALGVAILPGCCEAWHWIVG
NDDLIA*

Fig. 4-19

>ORF38640c (SEQ ID NO:409)
LTSLNIAPCERLKAVGSAIAGDVNITGSGSSVVQQCAGAPFLGFGRTGGSRLHLVTYKVARPVLVTGKAGLVAVADQLVA
GSVILQRGSHCWFGGGGVALGNPCQDLIESLNQRVR*

>ORF39309c (SEQ ID NO:411)
SCLVIIASIAVAGNYCPLSAVQSGYGRFGEGVTKVRVGTAIVFPKQLDIVRLGYIECEQAATSSTKRSDQNSVPSSIEKG
IADDRCSVDEILVGTALVADAGKLVKLFEYCCRDESELFHAGFDVETMQHAIVGSNINSRAALVCLCLECAVVAAGFDGV
GQISEVLGRSDHGWRRVHDVP*

>ORF38768 (SEQ ID NO:413)
GTSCTRLQPWSDRPNTSLIWPTPSNPAATTAHSRQRQTSAALEFMLDPTMACCMVSTSKPAWKSSLSSLQQYSKSLTSLP
ASATRAVPTNISSTLHRSSAMPFSMELGTLF*

>ORF40047c (SEQ ID NO:415)
KHLAGSEHIFDHHFAFLAAIDLQIQSPAIASHIRVRGTGCSVFVPLHRVLANKPNIPVHATECGVRLCCHGLGCEGVAVD
GWGTVCAFALGLAIPDTIERHGSAGIGLSVIVLEVFSCTYDDDVALTARMAYQGRRSRDGLTSIVVGASTEKRGAAERYS
ESRRIYISGLEGIDRGICVWVVVAGIADQIEAPDISLQIASIGVVSNAIAVVISQARRR*

>ORF40560c (SEQ ID NO:417)
PASPVVGHGWASDMSAHRLLPDRRQDEHPRPRPHSRTYSWHSRQRLGKPELELQELWREDVRGSCSFLLDDPPGLSSVTL
VAGVEDDALITSASLGLTLFVLGNGQAT*

>ORF40238 (SEQ ID NO:419)
VAWPLPRTNSVNPRLALVMSASSSTPATRVTDDKPGGSSRRNEHEPLTSSRHSSCSSSSGLPNLCLECHEYVRECGRGRG
CSSCRRSGSSRWADISLAQPCPTTGLAGHILGTSGTDSVFLRQTHQRPARNRVVLHYQAGPSRSLRIGAATMKSNRGFTL
IELMIVVVIIAILAGIAYPSYDEYVKRGNRTEGQALLSEAAATQERYFSQNNTYITTQADIGKLHMRNTSGTTVKSSTGK
YSLTVDTVANDGGYRLIANQAFNDLDCGNLTLTANGEKGRTGSKKSVAECWR*

>ORF40329 (SEQ ID NO:421)
RTTNLAGRRGEMNMNPLRLLATALAALALACPTFALSATNTFENVGVVEDVHPAAGLVVVDGQTYRLPNRVQQQDSPVIF
LVRQGQTVSFSGKLTSDLPEIESFYIIKQAPLVPFGSEQQQ*

Fig. 4-20

>ORF40709c (SEQ ID NO:423)
SLCSTSLLLLRSEGNERGLLDNVERLDFWQVAGEFAGERHCLSLTYQEYDRRVLLLDTVGQAICLPIDYYQTGGRMNILD
HAHILERIRGTQGKGWASQS*

>ORF40507 (SEQ ID NO:425)
SMGRHIACPTVSNNRTRRSYSWYVRDRQCLSPANSPATCQKSSRSTLSSRPLSFPSDRSSNNEVEQRLHSHRVDDRRSNH
RYSCWYRLPQLRRIREARESHRRTGITQRSSRYSRALFFTEQYLYHYPSRHRQAAYAQHIGHHSEVLHRQIQPYRRYGSQ
RRRLSPYR*

>ORF41275c (SEQ ID NO:427)
VGGVGRAGTGGTGGEHYLLDAQGTAGQGFVIGIFAHLKKPGLCRALPFSCLGALAPAFCNALLASSPAFLAVGGQGQVAT
IKIVECLVSDKAITSVVGYRIDGKAVFACGGLHCGARCVAHMQLADVGLGSDISIVL*

>ORF42234c (SEQ ID NO:429)
STSSRPEPSVAAPFPSGEGGSKVHSSNHRVAHETAMQIKLANPRGFCAGVDRAIEIVNRALDVFGPPIYVRHEVVHNKFV
VDNLRQRGAIFVEELDQVPDNVIVIFSAHGVSQAVRKEAEGRGLKVFDATCPLVTKVHMEVVRYSRDGHECVLIGHEGHP
EVEGTMGQYDASNGGAIYLVEDEADVAALEVRKPEALHYVTQTTLSMDDTSKVIDALRAKFPQIQGPRKNDICYATQNRQ
DAVKELADQCDMVLVVGSPNSSNSNRLRELAERMGTPAYLIDGAEDMQRGWFDGVRRIGITAGASAPEVLVRGVIAQLRE
WGASEEQELEGREENITFSMPKELRVKAL*

Fig. 4-21

>ORF41764c (SEQ ID NO:431)
RPPRGGRHHGPVRCQQRRCHLPGGGRGRRRRAGGAQARSPALRDPDHPVDGRHLEGHRCPARQVPADPGAAQERHLLCHP
EPPGCREGTGRPVRHGPGGGQPQQFQLQPPARTRRAHGHAGLPDRRRRGHATRLVRRCASHRNHRRRLRAGSAGARSDRP
AT*

>ORF41284 (SEQ ID NO:433)
LGDHSAHQHFRRGGACGDSDATHTVEPAALHVLGAVDQVGRRAHALGEFAQAVGVGTVGAAHHQDHVALVGQFLHGILAV
LGGIADVVLARPLDLRELGAQGIDDLRGVVHRQGGLGHVVQGFGLAHLQRGDVGLVLHQVDGTAVAGIVLAHGAFHLGVA
FMPDQHAFVAVAAVAHHFHVHLGHQRAGRVENLQAAPLGFLADRLGNAVGAEDDDDVVRHLIEFLDEDGAALAQVVHDEL
VVHHLVTHVDRRAEDIEGTVDDLDGAIHAGAEAAGIGEFDLHGGLVGDAVIGRMNLATALPAWEGRSDRRFRPAGRR

>ORF41598 (SEQ ID NO:435)
PSRCRPSTGWSGSRSAGLRACAPPARRRRPRPPPGRWHRRCWHRTGPWCLPPRGGLHARSARIRGRRGCSAPLPCAPWSP
AGRSRRKPSGRAPRLPCGPPGKRRGR*

>ORF42172c (SEQ ID NO:437)
QGSFVQSPRRPRDRHANQTRQSPRLLRRRGSRHRDRQPCPRCLRPADLRASRGGAQQVRRGQPAPARRHLRRGTRSGAGQ
RHRHLQRPRRFPGGPQGSRGARPEGFRRDLPAGDQGAHGSGALQPRRPRMRADRA*

>ORF42233c (SEQ ID NO:152)
RRPAGLNRRSLRPSQAGRAVARFIRPITASPTRPPCKSNSPIPAASAPAWIAPSRSSTVPSMSSARRSTCVTRWCTTSSS
WTTCASAAPSSSRNSIRCRTTSSSSSAPTAFPRRSARKPRGAA*

Fig. 4-22

33A9 (SEQ ID NO:102)

CAAAGCATAAGACCAAGATGGCACATTGCCAACAAAATACCCTTCCCCGCTACCGTTGTTTTATCGTTGTTGCCAGCCCT
GATCTGGCGGAAAAGCCCGCTCCATGAATCGTCATGGAGCCTCCCATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCC
GTCCCCCACCCCAACAACCAAAGCTGCCCCAGGGGGATTCATCCTTCCTCTGAGCAGCATGGAACTGCTCGGCACGCCTC
GCCGCCGGCAGCTACTGGAGAACATCTGGCAGCGCGCCTCGCTATCCAAGCAGCAATTCGAGGAGATCTACCGGCGGCCA
CTGGCCAACTATGCCGAGCTGGTCCAGCAGCTCCCTGCTTCGGAAAATCATCACCATGCCCATCCAGGCGGGATGATCGA
TCACGGCCTGGAGATCGTGGCCTACGCACTCAAGGTACGGCAGACCTACCTGCTCCCGATCGGCGCAGCGCCGGAGTCAC
AGTCAGCCCAGGCTGAAGCCTGGTCGGCCGCCGCGGCGTATGGCGCCCTGGCTCATGACATAGGCAAGATCGTCGTCGAC
CTGCAGGTTGAGCTACAGGACGGCAGCACCTGGCACCCTTGGAACGGACCGATCAACCAGCCATACCGCTTCAAGTACGT
GAAGTCCCGCGAATACCAGCTCCACGGCGCTGCCTCAGCACTTCTCATCCACCAACTGCTACCGCGCACTGCACTCGATT
GGCTCAGTCGCTTTCCAGAGCTGTGGGCTCAATTGATCTACCTGTTCGCTGGGCAGTACGAGCACGCCGGGATCCTCGGC
GAGATCATCGTGAAGGCAGACCAGGCCTCAGTTGCACAGGAGCTAGGAGGCAATCCGGATCGAGCTCTGGCTGCACCGAA
GCAGTCGCTGCAGCGGCAGTTGGCAGACGGCCTTCGCTTCTTGGTGAAGGACAAGTTCAAGTTGAATCAACCTAGCGGCC
CGTCTGATGGATGGCTGACCCAGGACGCACTCTGGCTGGTGAGCAAGCCTGCTGCCGATCAACTGAGAGCCTACCTGCTG
GCCCAGGGTATCGATGGGGTGCCCTCCTCTAACGCGCCGTTCTTCAGCATGCTCCAGGACCAAGCCGTCATCCAGACAAA
TGCCGAGGACAAGGCCATTTGGACGGCCACGGTAGACAACGGTGCTGGATGGAGAAACAAGTTCACGCTACTCAAGATTG
CTCCAGCCTTGATCTGGACAGATGCTGCCGAGCGCCCCTCACCCTACAGCGGATCACTGGTCGTTGAAGATGGAACCGCC
TCAACGGAAAAGCCGGAAACGACCTGTGAAATTCCCAACGGGCCGGCTGAACAGCAGCAAGCACCAGAAACGAAGATGAT
GCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGATGATCAAG
AAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACGACTCGCCG
GCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGATTGCGCTCC
TGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATCTGGCA
TCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCACGCCA
GGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGGAAGCT
GGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAAGGTTTCTG
GTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGGACAAC
CCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGACGCCGCAGCAGCTCACCGAGGAGTACATCTTCGCG
CACGATCTCCGGGAAGCCAGC

Fig. 5

33A9 SEQ ID NO:103

```
  1 MNRHGASHVS TPFLDIQEAV PHPNNQSCPR GIHPSSEQHG TARHASPPAA
 51 TGEHLAARLA IQAAIRGDLP AATGQLCGVG PAxPCFGKSS PCPSRRDDRS
101 RPGDRGYALK VRQTYLLPIG AAPESQSAQA EAWSAAAAYG ALAHDIGKIV
151 VDLQVELQDG STWHPWNGPI NQPYRFKYVK SREYQLHGAA SALFIHQLLP
201 RTALDWLSRF PELWAQLIYL FAGQYEHAGI LGEIIVKADQ ASVAQELGGN
251 PDRALAAPKQ SLQRQLADGL RFLVKDKFKL NQPSGPSDGW LTQDALWLVS
301 KPAADQLRAY LLAQGIDGVP SSNAPFFSML QDQAVIQTNA EDKAIWTATV
351 DNGAGWRNKF TLLKIAPALI WTDAAERPSP YSGSLVVEDG TASTEKPETT
401 CEIPNGPAEQ QQAPETKMML HQPAPSVAKP ANETQAIAKP STDDQEETDD
451 LYALLGNINS PLEELDTSHD SPAASPTNTR GEENLQQPLG TKEPTDCAPE
501 AIEDVFMPSR STDLGQGFVG WMKSGIAARR LFINDTKALV HTVDGTAMLV
551 TPGIFKRYVQ EHPVLEKLAQ AKETTGWKLV QRAFEKQGLH RKTSKNLNIW
601 TIKVSGPRKT KELKAYLLQD PKLLFPEQPL DNPSLTVITD AE*
```

Fig. 6A

33A9--ORF1 SEQ ID NO: 189

ATGGAGCCTCCCATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCCGTCCCCCACCCCAACAACCAAAGCTGCCCCAGG
GGGATTCATCCTTCCTCTGAGCAGCATGGAACTGCTCGGCACGCCTCGCCGCCGGCAGCTACTGGAGAACATCTGGCAGC
GCGCCTCGCTATCCAAGCAGCAATTCGAGGAGATCTACCGGCGGCCACTGGCCAACTATGCCGAGCTGGTCCAGCAGCTC
CCTGCTTCGGAAAATCATCACCATGCCCATCCAGGCGGGATGATCGATCACGGCCTGGAGATCGTGGCCTACGCACTCAA
GGTACGGCAGACCTACCTGCTCCCGATCGGCGCAGCGCCGGAGTCACAGTCAGCCCAGGCTGAAGCCTGGTCGGCCGCCG
CGGCGTATGGCGCCCTGGCTCATGACATAGGCAAGATCGTCGTCGACCTGCAGGTTGAGCTACAGGACGGCAGCACCTGG
CACCCTTGGAACGGACCGATCAACCAGCCATACCGCTTCAAGTACGTGAAGTCCCGCGAATACCAGCTCCACGGCGCTGC
CTCAGCACTTCTCATCCACCAACTGCTACCGCGCACTGCACTCGATTGGCTCAGTCGCTTTCCAGAGCTGTGGGCTCAAT
TGATCTACCTGTTCGCTGGGCAGTACGAGCACGCCGGGATCCTCGGCGAGATCATCGTGAAGGCAGACCAGGCCTCAGTT
GCACAGGAGCTAGGAGGCAATCCGGATCGAGCTCTGGCTGCACCGAAGCAGTCGCTGCAGCGGCAGTTGGCAGACGGCCT
TCGCTTCTTGGTGAAGGACAAGTTCAAGTTGAATCAACCTAGCGGCCCGTCTGATGGATGGCTGACCCAGGACGCACTCT
GGCTGGTGAGCAAGCCTGCTGCCGATCAACTGAGAGCCTACCTGCTGGCCCAGGGTATCGATGGGGTGCCCTCCTCTAAC
GCGCCGTTCTTCAGCATGCTCCAGGACCAAGCCGTCATCCAGACAAATGCCGAGGACAAGGCCATTTGGACGGCCACGGT
AGACAACGGTGCTGGATGGAGAAACAAGTTCACGCTACTCAAGATTGCTCCAGCCTTGATCTGGACAGATGCTGCCGAGC
GCCCCTCACCCTACAGCGGATCACTGGTCGTTGAAGATGGAACCGCCTCAACGGAAAAGCCGGAAACGACCTGTGAAATT
CCCAACGGGCCGGCTGAACAGCAGCAAGCACCAGAAACGAAGATGATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACC
GGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGATGATCAAGAAGAAACAGACGATTTGTATGCACTTCTTGGTA
AATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACGACTCGCCGGCTGCCTCTCCTACGAACACACGCGGGGAGGAG
AACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGATTGCGCTCCTGAAGCAATTGAAGATGTATTTATGCCTAGCAG
AAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATCTGGCATCGCGGCCCGTCGCCTGTTCATCAACGACACCA
AGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCG
GTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCA
TCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACC
TGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGA
GGTGTGGAATGA

Fig. 6B

33A9--ORF2 SEQ ID NO:190

ATGTTTCAACTCCTTTCCTGGATATCCAGGAAGCCGTCCCCCACCCCAACAACCAAAGCTGCCCCAGGGGGATTCATCCT
TCCTCTGAGCAGCATGGAACTGCTCGGCACGCCTCGCCGCCGGCAGCTACTGGAGAACATCTGGCAGCGCGCCTCGCTAT
CCAAGCAGCAATTCGAGGAGATCTACCGGCGGCCACTGGCCAACTATGCCGAGCTGGTCCAGCAGCTCCCTGCTTCGGAA
AATCATCACCATGCCCATCCAGGCGGGATGATCGATCACGGCCTGGAGATCGTGGCCTACGCACTCAAGGTACGGCAGAC
CTACCTGCTCCCGATCGGCGCAGCGCCGGAGTCACAGTCAGCCCAGGCTGAAGCCTGGTCGGCCGCCGCGGCGTATGGCG
CCCTGGCTCATGACATAGGCAAGATCGTCGTCGACCTGCAGGTTGAGCTACAGGACGGCAGCACCTGGCACCCTTGGAAC
GGACCGATCAACCAGCCATACCGCTTCAAGTACGTGAAGTCCCGCGAATACCAGCTCCACGGCGCTGCCTCAGCACTTCT
CATCCACCAACTGCTACCGCGCACTGCACTCGATTGGCTCAGTCGCTTTCCAGAGCTGTGGGCTCAATTGATCTACCTGT
TCGCTGGGCAGTACGAGCACGCCGGGATCCTCGGCGAGATCATCGTGAAGGCAGACCAGGCCTCAGTTGCACAGGAGCTA
GGAGGCAATCCGGATCGAGCTCTGGCTGCACCGAAGCAGTCGCTGCAGCGGCAGTTGGCAGACGGCCTTCGCTTCTTGGT
GAAGGACAAGTTCAAGTTGAATCAACCTAGCGGCCCGTCTGATGGATGGCTGACCCAGGACGCACTCTGGCTGGTGAGCA
AGCCTGCTGCCGATCAACTGAGAGCCTACCTGCTGGCCCAGGGTATCGATGGGGTGCCCTCCTCTAACGCGCCGTTCTTC
AGCATGCTCCAGGACCAAGCCGTCATCCAGACAAATGCCGAGGACAAGGCCATTTGGACGGCCACGGTAGACAACGGTGC
TGGATGGAGAAACAAGTTCACGCTACTCAAGATTGCTCCAGCCTTGATCTGGACAGATGCTGCCGAGCGCCCCTCACCCT
ACAGCGGATCACTGGTCGTTGAAGATGGAACCGCCTCAACGGAAAAGCCGGAAACGACCTGTGAAATTCCCAACGGGCCG
GCTGAACAGCAGCAAGCACCAGAAACGAAGATGATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGAC
GCAGGCGATTGCGAAACCCTCAACTGATGATCAAGAAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGC
CACTAGAAGAGCTAGACACTAGCCACGACTCGCCGGCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAG
CCACTAGGGACCAAGGAGCCAACAGATTGCGCTCCTGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCT
GGGACAGGGATTCGTTGGTTGGATGAAATCTGGCATCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGC
ATACCGTAGACGGGACCGCCATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAA
CTGGCCCAAGCCAAGGAGACGACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAG
TAAAAACCTGAACATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATC
CCAAATTGCTGTTCCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Fig. 6C

ORF 3— 33A9 SEQ ID NO:191

ATGGAACTGCTCGGCACGCCTCGCCGCCGGCAGCTACTGGAGAACATCTGGCAGCGCGCCTCGCTATCCAAGCAGCAATT
CGAGGAGATCTACCGGCGGCCACTGGCCAACTATGCCGAGCTGGTCCAGCAGCTCCCTGCTTCGGAAAATCATCACCATG
CCCATCCAGGCGGGATGATCGATCACGGCCTGGAGATCGTGGCCTACGCACTCAAGGTACGGCAGACCTACCTGCTCCCG
ATCGGCGCAGCGCCGGAGTCACAGTCAGCCCAGGCTGAAGCCTGGTCGGCCGCCGCGGCGTATGGCGCCCTGGCTCATGA
CATAGGCAAGATCGTCGTCGACCTGCAGGTTGAGCTACAGGACGGCAGCACCTGGCACCCTTGGAACGGACCGATCAACC
AGCCATACCGCTTCAAGTACGTGAAGTCCCGCGAATACCAGCTCCACGGCGCTGCCTCAGCACTTCTCATCCACCAACTG
CTACCGCGCACTGCACTCGATTGGCTCAGTCGCTTTCCAGAGCTGTGGGCTCAATTGATCTACCTGTTCGCTGGGCAGTA
CGAGCACGCCGGGATCCTCGGCGAGATCATCGTGAAGGCAGACCAGGCCTCAGTTGCACAGGAGCTAGGAGGCAATCCGG
ATCGAGCTCTGGCTGCACCGAAGCAGTCGCTGCAGCGGCAGTTGGCAGACGGCCTTCGCTTCTTGGTGAAGGACAAGTTC
AAGTTGAATCAACCTAGCGGCCCGTCTGATGGATGGCTGACCCAGGACGCACTCTGGCTGGTGAGCAAGCCTGCTGCCGA
TCAACTGAGAGCCTACCTGCTGGCCCAGGGTATCGATGGGGTGCCCTCCTCTAACGCGCCGTTCTTCAGCATGCTCCAGG
ACCAAGCCGTCATCCAGACAAATGCCGAGGACAAGGCCATTTGGACGGCCACGGTAGACAACGGTGCTGGATGGAGAAAC
AAGTTCACGCTACTCAAGATTGCTCCAGCCTTGATCTGGACAGATGCTGCCGAGCGCCCCTCACCCTACAGCGGATCACT
GGTCGTTGAAGATGGAACCGCCTCAACGGAAAAGCCGGAAACGACCTGTGAAATTCCCAACGGGCCGGCTGAACAGCAGC
AAGCACCAGAAACGAAGATGATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCAGGCGATTGCG
AAACCCTCAACTGATGATCAAGAAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCACTAGAAGAGCT
AGACACTAGCCACGACTCGCCGGCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCACTAGGGACCA
AGGAGCCAACAGATTGCGCTCCTGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTC
GTTGGTTGGATGAAATCTGGCATCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGG
GACCGCCATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCA
AGGAGACGACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAAC
ATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTT
CCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Fig. 6D

ORF 4--33A9 SEQ ID NO:192

```
ATGATCGATCACGGCCTGGAGATCGTGGCCTACGCACTCAAGGTACGGCAGACCTACCTGCTCCCGATCGGCGCAGCGCC
GGAGTCACAGTCAGCCCAGGCTGAAGCCTGGTCGGCCGCCGCGGCGTATGGCGCCCTGGCTCATGACATAGGCAAGATCG
TCGTCGACCTGCAGGTTGAGCTACAGGACGGCAGCACCTGGCACCCTTGGAACGGACCGATCAACCAGCCATACCGCTTC
AAGTACGTGAAGTCCCGCGAATACCAGCTCCACGGCGCTGCCTCAGCACTTCTCATCCACCAACTGCTACCGCGCACTGC
ACTCGATTGGCTCAGTCGCTTTCCAGAGCTGTGGGCTCAATTGATCTACCTGTTCGCTGGGCAGTACGAGCACGCCGGGA
TCCTCGGCGAGATCATCGTGAAGGCAGACCAGGCCTCAGTTGCACAGGAGCTAGGAGGCAATCCGGATCGAGCTCTGGCT
GCACCGAAGCAGTCGCTGCAGCGGCAGTTGGCAGACGGCCTTCGCTTCTTGGTGAAGGACAAGTTCAAGTTGAATCAACC
TAGCGGCCCGTCTGATGGATGGCTGACCCAGGACGCACTCTGGCTGGTGAGCAAGCCTGCTGCCGATCAACTGAGAGCCT
ACCTGCTGGCCCAGGGTATCGATGGGGTGCCCTCCTCTAACGCGCCGTTCTTCAGCATGCTCCAGGACCAAGCCGTCATC
CAGACAAATGCCGAGGACAAGGCCATTTGGACGGCCACGGTAGACAACGGTGCTGGATGGAGAAACAAGTTCACGCTACT
CAAGATTGCTCCAGCCTTGATCTGGACAGATGCTGCCGAGCGCCCCTCACCCTACAGCGGATCACTGGTCGTTGAAGATG
GAACCGCCTCAACGGAAAAGCCGGAAACGACCTGTGAAATTCCCAACGGGCCGGCTGAACAGCAGCAAGCACCAGAAACG
AAGATGATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGA
TGATCAAGAAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACG
ACTCGCCGGCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGAT
TGCGCTCCTGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAA
ATCTGGCATCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGG
TCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGC
TGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAA
GGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTC
TGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA
```

Fig. 6E

ORF5--33A9 SEQ ID NO:193

```
ATGCTCCAGGACCAAGCCGTCATCCAGACAAATGCCGAGGACAAGGCCATTTGGACGGCCACGGTAGACAACGGTGCTGG
ATGGAGAAACAAGTTCACGCTACTCAAGATTGCTCCAGCCTTGATCTGGACAGATGCTGCCGAGCGCCCCTCACCCTACA
GCGGATCACTGGTCGTTGAAGATGGAACCGCCTCAACGGAAAAGCCGGAAACGACCTGTGAAATTCCCAACGGGCCGGCT
GAACAGCAGCAAGCACCAGAAACGAAGATGATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCA
GGCGATTGCGAAACCCTCAACTGATGATCAAGAAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCAC
TAGAAGAGCTAGACACTAGCCACGACTCGCCGGCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCA
CTAGGGACCAAGGAGCCAACAGATTGCGCTCCTGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGG
ACAGGGATTCGTTGGTTGGATGAAATCTGGCATCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATA
CCGTAGACGGGACCGCCATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTG
GCCCAAGCCAAGGAGACGACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAA
AAACCTGAACATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCA
AATTGCTGTTCCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA
```

Fig. 6F

ORF6--33A9 SEQ ID NO:194

ATGATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGATGA
TCAAGAAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACGACT
CGCCGGCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGATTGC
GCTCCTGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATC
TGGCATCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCA
CGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGG
AAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAAGGT
TTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGG
ACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Fig. 6G

ORF7--33A9 SEQ ID NO:195

ATGCTCCATCAACCTGCGCCGAGCGTTGCGAAACCGGCAAACGAGACGCAGGCGATTGCGAAACCCTCAACTGATGATCA
AGAAGAAACAGACGATTTGTATGCACTTCTTGGTAATATCAATTCGCCACTAGAAGAGCTAGACACTAGCCACGACTCGC
CGGCTGCCTCTCCTACGAACACACGCGGGGAGGAGAACCTACAGCAGCCACTAGGGACCAAGGAGCCAACAGATTGCGCT
CCTGAAGCAATTGAAGATGTATTTATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATCTGG
CATCGCGGCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCACGC
CAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGGAAG
CTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAAGGTTTC
TGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGGACA
ACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Fig. 6H

ORF8--33A9 SEQ ID NO:196

ATGCCTAGCAGAAGTACTGATCTGGGACAGGGATTCGTTGGTTGGATGAAATCTGGCATCGCGGCCCGTCGCCTGTTCAT
CAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCC
AAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAA
CAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCT
CAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCG
ATGCCGAAGGAGGTGTGGAATGA

Fig. 6I

ORF9--33A9 SEQ ID NO:197

ATGAAATCTGGCATCGCGGCCCCGTCGCCTGTTCATCAACGACACCAAGGCTTTGGTGCATACCGTAGACGGGACCGCCAT
GCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGACGA
CCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGACC
ATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAGCA
GCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Fig. 6J

ORF10-33A9 SEQ ID NO:198

ATGCTGGTCACGCCAGGAATTTTCAAGCGCTATGTCCAAGAGCATCCGGTGCTTGAAAAACTGGCCCAAGCCAAGGAGAC
GACCGGCTGGAAGCTGGTGCAGCGCGCGTTCGAAAAACAGGGGCTTCATCGGAAGACCAGTAAAAACCTGAACATCTGGA
CCATCAAGGTTTCTGGTCCTCGCAAGACGAAAGAGCTCAAGGCCTACCTGCTCCAGGATCCCAAATTGCTGTTCCCTGAG
CAGCCTCTGGACAACCCAAGCCTCACGGTCATCACCGATGCCGAAGGAGGTGTGGAATGA

Fig. 6K

ORF 1--33A9 SEQ ID NO:199

MEPPMFQLLSWISRKPSPTPTTKAAPGGFILPLSSMELLGTPRRRQLLENIWQRASLSKQQFEEIYRRPLANYAELVQQL
PASENHHHAHPGGMIDHGLEIVAYALKVRQTYLLPIGAAPESQSAQAEAWSAAAAYGALAHDIGKIVVDLQVELQDGSTW
HPWNGPINQPYRFKYVKSREYQLHGAASALLIHQLLPRTALDWLSRFPELWAQLIYLFAGQYEHAGILGEIIVKADQASV
AQELGGNPDRALAAPKQSLQRQLADGLRFLVKDKFKLNQPSGPSDGWLTQDALWLVSKPAADQLRAYLLAQGIDGVPSSN
APFFSMLQDQAVIQTNAEDKAIWTATVDNGAGWRNKFTLLKIAPALIWTDAAERPSPYSGSLVVEDGTASTEKPETTCEI
PNGPAEQQQAPETKMMLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEE
NLQQPLGTKEPTDCAPEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHP
VLEKLAQAKETTGWKLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEG
GVE.

Fig. 6L

ORF2--33A9 SEQ ID NO:200

MFQLLSWISRKPSPTPTTKAAPGGFILPLSSMELLGTPRRRQLLENIWQRASLSKQQFEEIYRRPLANYAELVQQLPASE
NHHHAHPGGMIDHGLEIVAYALKVRQTYLLPIGAAPESQSAQAEAWSAAAAYGALAHDIGKIVVDLQVELQDGSTWHPWN
GPINQPYRFKYVKSREYQLHGAASALLIHQLLPRTALDWLSRFPELWAQLIYLFAGQYEHAGILGEIIVKADQASVAQEL
GGNPDRALAAPKQSLQRQLADGLRFLVKDKFKLNQPSGPSDGWLTQDALWLVSKPAADQLRAYLLAQGIDGVPSSNAPFF
SMLQDQAVIQTNAEDKAIWTATVDNGAGWRNKFTLLKIAPALIWTDAAERPSPYSGSLVVEDGTASTEKPETTCEIPNGP
AEQQQAPETKMMLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEENLQQ
PLGTKEPTDCAPEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEK
LAQAKETTGWKLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6M

ORF2-33A9 SEQ ID NO:201

MELLGTPRRRQLLENIWQRASLSKQQFEEIYRRPLANYAELVQQLPASENHHHAHPGGMIDHGLEIVAYALKVRQTYLLP
IGAAPESQSAQAEAWSAAAAYGALAHDIGKIVVDLQVELQDGSTWHPWNGPINQPYRFKYVKSREYQLHGAASALLIHQL
LPRTALDWLSRFPELWAQLIYLFAGQYEHAGILGEIIVKADQASVAQELGGNPDRALAAPKQSLQRQLADGLRFLVKDKF
KLNQPSGPSDGWLTQDALWLVSKPAADQLRAYLLAQGIDGVPSSNAPFFSMLQDQAVIQTNAEDKAIWTATVDNGAGWRN
KFTLLKIAPALIWTDAAERPSPYSGSLVVEDGTASTEKPETTCEIPNGPAEQQQAPETKMMLHQPAPSVAKPANETQAIA
KPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEENLQQPLGTKEPTDCAPEAIEDVFMPSRSTDLGQGF
VGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKLAQAKETTGWKLVQRAFEKQGLHRKTSKNLN
IWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6N

ORF4-33A9 SEQ ID NO:202

MIDHGLEIVAYALKVRQTYLLPIGAAPESQSAQAEAWSAAAAYGALAHDIGKIVVDLQVELQDGSTWHPWNGPINQPYRF
KYVKSREYQLHGAASALLIHQLLPRTALDWLSRFPELWAQLIYLFAGQYEHAGILGEIIVKADQASVAQELGGNPDRALA
APKQSLQRQLADGLRFLVKDKFKLNQPSGPSDGWLTQDALWLVSKPAADQLRAYLLAQGIDGVPSSNAPFFSMLQDQAVI
QTNAEDKAIWTATVDNGAGWRNKFTLLKIAPALIWTDAAERPSPYSGSLVVEDGTASTEKPETTCEIPNGPAEQQQAPET
KMMLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEENLQQPLGTKEPTD
CAPEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKLAQAKETTG
WKLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6O

ORF5--33A9 SEQ ID NO:203

MLQDQAVIQTNAEDKAIWTATVDNGAGWRNKFTLLKIAPALIWTDAAERPSPYSGSLVVEDGTASTEKPETTCEIPNGPA
EQQQAPETKMMLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEENLQQP
LGTKEPTDCAPEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKL
AQAKETTGWKLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6P

ORF6-33A9 SEQ ID NO:204

MMLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEENLQQPLGTKEPTDC
APEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKLAQAKETTGW
KLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6Q

ORF7-33A9 SEQ ID NO:205

MLHQPAPSVAKPANETQAIAKPSTDDQEETDDLYALLGNINSPLEELDTSHDSPAASPTNTRGEENLQQPLGTKEPTDCA
PEAIEDVFMPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKLAQAKETTGWK
LVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6R

ORF8--33A9 SEQ ID NO:206

MPSRSTDLGQGFVGWMKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKLAQAKETTGWKLVQRAFEK
QGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6S

ORF9--33A9 SEQ ID NO:207

MKSGIAARRLFINDTKALVHTVDGTAMLVTPGIFKRYVQEHPVLEKLAQAKETTGWKLVQRAFEKQGLHRKTSKNLNIWT
IKVSGPRKTKELKAYLLQDPKLLFPEQPLDNPSLTVITDAEGGVE.

Fig. 6T

ORF10--33A9 SEQ ID NO:208

MLVTPGIFKRYVQEHPVLEKLAQAKETTGWKLVQRAFEKQGLHRKTSKNLNIWTIKVSGPRKTKELKAYLLQDPKLLFPE
QPLDNPSLTVITDAEGGVE.

Fig. 6U

34B12 EcoR1 fragment Map (1 > 4590) ORF Map
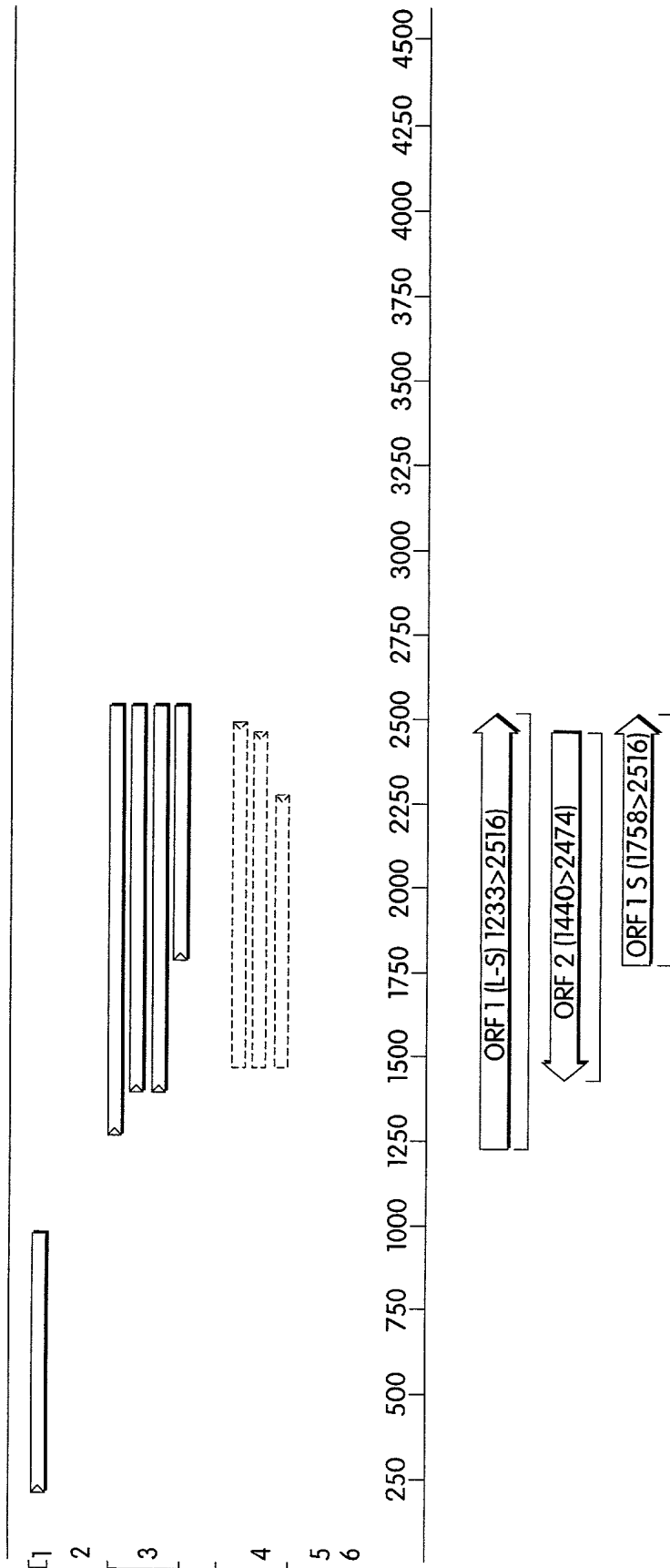
Fig. 7A

Sequence: 34B12 EcoR1 fragment From: 1 To: 4590

```
                10         20         30         40         50         60
         |↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|
GAATTCCATG GCGCCGTGGA GGAGGCTTCC GAGTCGCCGG TGGCAGGCGT ACGGGCCGGC 60
AACTACCAGG TCGACCTGGA CGATGCGAGC TTTGCCCGCC AGGTAGAACG CCTGCAGGCC 120
CACGTGAGGG CCGGCGACGT GTTCCAGATC GTACCTTCGC GCAGCTTCAG CATGCCGTGC 180
GCGGACCCCT GGCGGGCCTA TCGCCAGTTG TGCCTGCGCA ACCCCAGCCC GTACCGCTTC 240
TTCCTCGATG CGGGGGACTT CTGCCTGTTC GGCGCTTCGC CGGAGTCGGC ATTGAAGTAC 300
                310        320        330        340        350        360
         |↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|
GACGCGGAGA GTCGCGAGGT GGAACTCTAT CCCATTGCCG GCACCCGCCC GCGCGGATGC 360
GATGCCCGGG GCGCCATCGA TGCGGAACTG GACAATCGCC TGGAAGCGGA GTTGCGCCTG 420
GATGCCAAGG AGATCGCCGA GCACATGATG CTGGTCGACC TGGCGCGCAA CGATCTGGCG 480
CGCGTCTGCC GCAGCGGTAC CCGGCAGGTG CGCGACATGC TCAAGGTCGA TCGCTACAGC 540
CACGTGATGC ACCTGGTCTC GCGCGTGGCT GGCGAACTGC ACGGCGAACT GGATGCGCTG 600
                610        620        630        640        650        660
         |↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|
CATGCCTACC GTGCCTGCCT GAACATGGGC ACCCTGGTCG GCGCGCCGAA GGTCCGTGCC 660
ATGCAGTTGC TGCGGCAGTA CGAGGATGGC TATCGCGGCA GCTACGGTGG TGCGATCGGC 720
ATTCTCGACA GCGCCGGCAA CCTCGATACC AGCATTGTCA TCCGCTCCGC CGAGGTCCGC 780
GAAGGTATCG CGCGGGTTCG GGCAGGCGCC GGCGTGGTGC TGGATTCGGA TCCACGGCTG 840
GAGGCCGAGG AAACCCGCAA CAAGGCGCTG GCGGTGCTGA CCGCCGTGGC CGCTGCCGAA 900
                910        920        930        940        950        960
         |↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|
CGCGAAAGGG GAGAGCGCGA TGCGCATCAC GCTGTTGGAT AACTTCGATT CCTTCACCTA 960
CAACCTGGTC GAGCAGTTCT GCCTGCTCGG CGCGGAGGTC CGGGTGATGC GCAACGATAC 1020
GCCGTTGCCG ACGATCCAGG CGGCATTGCT GGCCGACGGT TGCGAACTGC TGGTGCTGTC 1080
GCCGGGGCCC GGTCGGCCGG AAGACGCCGG CTGTATGCTG GAATTGCTCG CCTGGGCCCG 1140
CGGGCGCTTG CCGGTGCTCG GCGTCTGCCT CGGCCACCAG GCGCTGGCGC TGGCCGCCGG 1200
                1210       1220       1230       1240       1250       1260
         |↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|  ↓↓↓↓↓|↓↓↓↓↓|
TGGCGCGGTG GGCGAGGCGA GGAAGCCGCT GCATGGCAAG AGCACGTCCC TGCGTTTCGA 1260
TCAGCGTCAC CCGCTGTTCG ACGGCATCGC TGACCTGCGC GTCGCGCGCT ACCACTCGCT 1320
GGTGGTCAGT CGCCTGCCGG AAGGTTTCGA CTGCCTGGCC GATGCCGATG GCGAGATCAT 1380
GGCGATGGCC GATCCGCGCA ATCGACAGCT GGGCTTGCAA TTCCATCCCG AGTCGATTCT 1440
CACCACCCAC GGCCAGCGTC TGCTGGAGAA CGCTCTACTC TGGTGCGGCG CGTTGGCGGT 1500
```

Fig. 7B

Sequence: 34B12 EcoRl fragment From: 1 To: 4590

```
          1510        1520        1530        1540        1550        1560
     լ ι ι ι ւ ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι
CGCGGAGCGC CTTCGGGCCT GAGCGGCGCT GCGCAGTTTC GACCGAGGCT CGGTTGCCAG 1560
GCCGGCGCAT CGTCGAAACG CTGGCGGCCC AGTTCGCGCA GGCGCTGGCG GGCGCTTTCG 1620
AGAAAGCGAC GGAAGCTGCG CTCGGATTCC AGCGCGGTGT TGTAGTAGCA ATACACCTTG 1680
GTGTCGATGC CGCCCGGTTC GTACAGTTCG CTGAGGACTG CCAGGGTACC GTTGCGCAGG 1740
CGTTCCTCGA CGAAATAATG CGGCGAGATG CCCCATCCGA CGCCGGCTTC CACCAGACGC 1800
          1810        1820        1830        1840        1850        1860
    լ ι ι ι ւ ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι
AGCATGTCGT CGAAGTTTTC CACGAAGAGC ACCTTGTCGC TGACCGGCCG CAGCAGGTTC 1860
GAATGCTGCC CGGAGCGGCT GCCGAGGCTG ATCTGCCGGT AATTGGCCAG GCTCGCGATG 1920
CTGTGCAGGG AGGCATTGCA CAACGGGTGC TGCGGATGGG CGACGACGAA CGCCTTGGTG 1980
TAGCCGAGCA CGCACTGGTT GAAGCGGGAG ATCTTCAGTT CCTCGTCGAT GGTGATGGCG 2040
ATATCGATTT CCGCGTTGTC CTGCTTGATC GTCGCCAGGC TATCGGCGGG CGAGGTGCGT 2100
          2110        2120        2130        2140        2150        2160
    լ ι ι ι ւ ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι
ATCAGGCTGA CCATGTTGAA ATCGTCGAGC AGTACGCTGC TCACCGTATC GCAGAACGAC 2160
GGCGGGATGG CGGTGTCCAG CAGCACCCGG AGATTGCGCG GACCCTTGTT GAGATTGAAG 2220
GCGATGTCGC CGATCAGCTG CTGGTAGTTC AGCAGGCTGC GCATGTAAGG GATCAGGCGA 2280
AGCGCCTGCT CGGTGGGTTC GACCTTGTAG CCGTCCCGAC GGACCAGCTC CACGCACAGG 2340
TCGATTTCCA GGTTGCTGAC CGCCGAGCTG ACCGCGGTGT GCGACTTGCG CAGGATCCGC 2400
          2410        2420        2430        2440        2450        2460
    լ ι ι ι ւ ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι
GCAGCGGAGG AAATCGAACC GGAGGCGATG ACCTGGAGGA ACATGTTCAC GTGATTCAGG 2460
TTATGAATAG GCATCCCTTA TTCCTTTTAT TGGGTGGCGC GTGCCGCTTC CCTTGATCGG 2520
GTCAGGTTGC CGCTACTGTG GAAGAAGCGT CGAGGACTCG ATAGATAGCG CCCGAGTGTT 2580
TCAACTTGTC TTCTGGATGA CGTTTTCATC GGGGAAACCT CCCGTCGGTC AGTGAATCGC 2640
AAGGGCTGGC GTGCGAGGGT GGAATCGGCC GCCGGCTCGC TTTCTGCGCG GCGGGCGCAC 2700
          2710        2720        2730        2740        2750        2760
    լ ι ι ι ւ ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι Ι ι ι ι ι ι Ι ι ι ι ι ι Ι
GGCACGGGGA GTCGTCGTTT TGGAGGTGAG GGATGACGGC TCTGTTTCAG GATATTTTTA 2760
TAATTATGTG AAAGAAGAGC TTATTTCAAC GAAATATGTT TCATATTGCT CGTTAAATTC 2820
GACGAAAAGA AAATCCGGAT ATTTACCGGT TATTTAACGC TAATACCAAG TGCCTAATAC 2880
CAAAGTATTA ACGCTGGTAT GCCGGCATGT CGTGTTCGGT CGTGGAGCGA GCCGAGCTAG 2940
GGACGGTTCT AATAAACCAA AAAATTATGT CGCGTACGTC TAACGACCGA AACCTATGTC 3000
```

Fig. 7C

Sequence: 34B12 EcoR1 fragment From: 1 To: 4590

```
          3010       3020       3030       3040       3050       3060
      |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
TCTTGTTAGC GTAGCCACCG GCCAGGCCGG TACGGACCCG GGATGGCCCT GGCGCGACCT 3060
ATGCGGTTAG AATCCGCGGC CTTGCAGGCG GATACCCGAG CTTCGCTCGA AGGTGTCGCG 3120
GTGCCGTGCC GTGGAATCGG CCGCCGGCTC GCTTTCTGCG CGGCGGGCGC ACGGCGACGG 3180
GGAGTCGTCG TTTTGGAGGT GAGGGATGAC GGCTCTGTTT CAGGATATTT TTATAATTAT 3240
GTGAAAGAAG AGCTTATTTC AACGAAATAT GTTTCATATT GCTCGTAAAT TCGACGAAAA 3300
          3310       3320       3330       3340       3350       3360
      |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
GAAAATCCGG ATATTTACCG GTTATTTAAC GTTAATACCA AGGGCCTAAT ACCAAAGTAT 3360
TAACGCTGGC ATGCCGGCAT GTCGTGTTCG GTCGTGGAGC GAGCCGAGCC AGGAACGGTT 3420
CTAAGAAACG AAAAAATTAT GTCGCGTAGG TCTAACGACC GAAACCTATG TCTTTTGTTA 3480
GCGTAGCCAC CGGCCAGGCC GGTACGGATG CCGGGATGGC CCTGGCGCGA CCTATGCGGT 3540
TAGAATCCGC GGCCTTGCAG GCGGATCCCC GGGGTTTGCT CAAGGGGACA CGGGTGCCGT 3600
          3610       3620       3630       3640       3650       3660
      |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
GCCCGAAACC TGCAATCGTC AGTTCCCTGC GGTCCAGCCT GCCGCCGGGT ATAAAATCGA 3660
GAGACGCGCT GTTGCGCCTT CAGGTGTAGC GACTATGACG CACATTTCCG AACGACTCCT 3720
GGTACAGGCC CACCTGGCCG CCAAGCAACC CCGTGTGTTG AGCGAGCAGG AGAGCGCCGA 3780
GCATCGCGCG GCGATCGCGG CCGAACTGAA GGCGCAAAAT GCTGTACTGG TGGCGCATTA 3840
CTACTGCGAC CCGGTGATCC AGGCGTTGGC CGAGGAGACC GGCGGTTGCG TATCCGATTC 3900
          3910       3920       3930       3940       3950       3960
      |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
GCTGGAGATG GCCCGTTTCG GCAACCAGCA TCCGGCGCAG ACGGTGGTCG TGGCCGGGGT 3960
GCGCTTCATG GGCGAGACGG CGAAGATCCT CAACCCTGAG AAGCGTGTGC TGATGCCGAC 4020
CCTCGAAGCG ACCTGCTCGC TCGACCTGGG ATGCCCGGTG GATGAATTCT CGGCTTTCTG 4080
CGACCAGCAC CCGGAACGGA CCGTGGTGGT CTATGCGAAC ACCTCCGCGG CGGTGAAGGC 4140
ACGCGCCGAC TGGGTCGTGA CCTCCAGTTG CGCGGTGGAG ATCGTCGAAC ACCTGATGGA 4200
          4210       4220       4230       4240       4250       4260
      |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
CAACGGCGAG CCCATCCTCT GGGCGCCGGA CCAGCACCTG GGACGCTACA TCCAGCGCGA 4260
GACGGGGGCC GACATGCTGC TCTGGGATGG CGCCTGTATC GTCCACGAGG AGTTCAAGGC 4320
CAAGCAGCTG GAAGACATGA AGGCGCTCTA CCCGGACGCC GCCATCCTGG TCCACCCCGA 4380
ATCGCCGGAA AGCGTGGTCG CGCTGGCCGA TGCCGTGGGC TCGACCAGCC AGTTGATCAA 4440
GGCCGCGCAG ACCCTGCCGA ACAAGACCTT CATCGTCGCC ACCGATCGCG GCATCTTCTA 4500
```

Fig. 7D

Sequence: 34B12 EcoR1 fragment From: 1 To: 4590

```
              4510        4520        4530        4540        4550        4560
    ┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴┴
CAAGATGCAG CAGTTGTGCC CGGACAAGGA TTTCATCGAG GCCCCCACCG CCGGCAACGG 4560
CGCCGCCTGC CGCAGTGCGC GCACTGCCCG 4590
```

Fig. 7E

Sequence: 34B12 ORF 1 L-S From: 1 To: 1284

```
            10          20          30          40          50          60
   |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
ATGGCAAGAG CACGTCCCTG CGTTTCGATC AGCGTCACCC GCTGTTCGAC GGCATCGCTG 60
ACCTGCGCGT CGCGCGCTAC CACTCGCTGG TGGTCAGTCG CCTGCCGGAA GGTTTCGACT 120
GCCTGGCCGA TGCCGATGGC GAGATCATGG CGATGGCCGA TCCGCGCAAT CGACAGCTGG 180
GCTTGCAATT CCATCCCGAG TCGATTCTCA CCACCCACGG CCAGCGTCTG CTGGAGAACG 240
CTCTACTCTG GTGCGGCGCG TTGGCGGTCG CGGAGCGCCT TCGGGCCTGA GCGGCGCTGC 300
           310         320         330         340         350         360
   |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
GCAGTTTCGA CCGAGGCTCG GTTGCCAGGC CGGCGCATCG TCGAAACGCT GGCGGCCCAG 360
TTCGCGCAGG CGCTGGCGGG CGCTTTCGAG AAAGCGACGG AAGCTGCGCT CGGATTCCAG 420
CGCGGTGTTG TAGTAGCAAT ACACCTTGGT GTCGATGCCG CCCGGTTCGT ACAGTTCGCT 480
GAGGACTGCC AGGGTACCGT TGCGCAGGCG TTCCTCGACG AAATAATGCG GCGAGATGCC 540
CCATCCGACG CCGGCTTCCA CCAGACGCAG CATGTCGTCG AAGTTTTCCA CGAAGAGCAC 600
           610         620         630         640         650         660
   |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
CTTGTCGCTG ACCGGCCGCA GCAGGTTCGA ATGCTGCCCG GAGCGGCTGC CGAGGCTGAT 660
CTGCCGGTAA TTGGCCAGGC TCGCGATGCT GTGCAGGGAG GCATTGCACA ACGGGTGCTG 720
CGGATGGGCG ACGACGAACG CCTTGGTGTA GCCGAGCACG CACTGGTTGA GCGGGAGAT 780
CTTCAGTTCC TCGTCGATGG TGATGGCGAT ATCGATTTCC GCGTTGTCCT GCTTGATCGT 840
CGCCAGGCTA TCGGCGGGCG AGGTGCGTAT CAGGCTGACC ATGTTGAAAT CGTCGAGCAG 900
           910         920         930         940         950         960
   |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
TACGCTGCTC ACCGTATCGC AGAACGACGG CGGGATGGCG GTGTCCAGCA GCACCCGGAG 960
ATTGCGCGGA CCCTTGTTGA GATTGAAGGC GATGTCGCCG ATCAGCTGCT GGTAGTTCAG 1020
CAGGCTGCGC ATGTAAGGGA TCAGGCGAAG CGCCTGCTCG GTGGGTTCGA CCTTGTAGCC 1080
GTCCCGACGG ACCAGCTCCA CGCACAGGTC GATTTCCAGG TTGCTGACCG CCGAGCTGAC 1140
CGCGGTGTGC GACTTGCGCA GGATCCGCGC AGCGGAGGAA ATCGAACCGG AGGCGATGAC 1200
           1210        1220        1230        1240        1250        1260
   |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
CTGGAGGAAC ATGTTCACGT GATTCAGGTT ATGAATAGGC ATCCCTTATT CCTTTTATTG
GGTGGCGCGT GCCGCTTCCC TTGA 1284
```

Fig. 7F

Sequence: 34B12 ORF 1 (L-S) PROTEIN From: 1 To: 427

```
              10          20          30          40          50
    |ıııı|ıııı| |ıııı|ıııı| |ıııı|ıııı| |ıııı|ıııı| |ıııı|ıııı|
MARARPCVSI SVTRCSTASL TCASRATTRW WSVACRKVST AWPMPMARSW
RWPIRAIDSW ACNSIPSRFS PPTASVCWRT LYSGAARWRS RSAFGPERRC
AVSTEARLPG RRIVETLAAQ FAQALAGAFE KATEAALGFQ RGVVVAIHLG
VDAARFVQFA EDCQGTVAQA FLDEIMRRDA PSDAGFHQTQ HVVEVFHEEY
LVADRPQQVR MLPGAATEAD LPVIGQARDA VQGGIAQRVL RMGDDERLGV
             260         270         280         290         300
    |ıııı|ıııı| |ıııı|ıııı| |ıııı|ıııı| |ıııı|ıııı| |ıııı|ıııı|
AEHALVEAGD LQFLVDGDGD IDFRVVLLDR RQAIGGRGAY QADHVEIVEQ
YAAHRIAERR RDGGVQQHPE IARTLVEIEG DVADQLLVVQ QAAHVRDQAK
RLLGGFDLVA VPTDQLHAQV DFQVADRRAD RGVRLAQDPR SGGNRTGGDD
LEEHVHVIQV MNRHPLFLLL GGACRFP 427
```

Fig. 7G

Sequence: 34B12 ORF 2 From: 1 To: 1035

```
              10        20        30        40        50        60
    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
ATGCCTATTC ATAACCTGAA TCACGTGAAC ATGTTCCTCC AGGTCATCGC CTCCGGTTCG  60
ATTTCCTCCG CTGCGCGGAT CCTGCGCAAG TCGCACACCG CGGTCAGCTC GGCGGTCAGC 120
AACCTGGAAA TCGACCTGTG CGTGGAGCTG GTCCGTCGGG ACGGCTACAA GGTCGAACCC 180
ACCGAGCAGG CGCTTCGCCT GATCCCTTAC ATGCGCAGCC TGCTGAACTA CCAGCAGCTG 240
ATCGGCGACA TCGCCTTCAA TCTCAACAAG GGTCCGCGCA ATCTCCGGGT GCTGCTGGAC 300
              310       320       330       340       350       360
    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
ACCGCCATCC CGCCGTCGTT CTGCGATACG GTGAGCAGCG TACTGCTCGA CGATTTCAAC 360
ATGGTCAGCC TGATACGCAC CTCGCCCGCC GATAGCCTGG CGACGATCAA GCAGGACAAC 420
GCGGAAATCG ATATCGCCAT CACCATCGAC GAGGAACTGA AGATCTCCCG CTTCAACCAG 480
TGCGTGCTCG GCTACACCAA GGCGTTCGTC GTCGCCCATC CGCAGCACCC GTTGTGCAAT 540
GCCTCCCTGC ACAGCATCGC GAGCCTGGCC AATTACCGGC AGATCAGCCT CGGCAGCCGC 600
              610       620       630       640       650       660
    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
TCCGGGCAGC ATTCGAACCT GCTGCGGCCG GTCAGCGACA AGGTGCTCTT CGTGGAAAAC 660
TTCGACGACA TGCTGCGTCT GGTGGAAGCC GGCGTCGGAT GGGGCATCTC GCCGCATTAT 720
TTCGTCGAGG AACGCCTGCG CAACGGTACC CTGGCAGTCC TCAGCGAACT GTACGAACCG 780
GGCGGCATCG ACACCAAGGT GTATTGCTAC TACAACACCG CGCTGGAATC CGAGCGCAGC 840
TTCCGTCGCT TTCTCGAAAG CGCCCGCCAG CGCCTGCGCG AACTGGGCCG CCAGCGTTTC 900
              910       920       930       940       950       960
    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
GACGATGCGC CGGCCTGGCA ACCGAGCCTC GGTCGAAACT GCGCAGCGCC GCTCAGGCCC 960
GAAGGCGCTC CGCGACCGCC AACGCGCCGC ACCAGAGTAG AGCGTTCTCC AGCAGACGCT 1020
GGCCGTGGGT GGTGA 1035
```

Fig. 7H

Sequence: 34B12 ORF 2 PROTEIN From:: 1 To: 344

```
              10         20         30         40         50         60
      |||||||||||  |||||||||||  |||||||||||  |||||||||||  |||||||||||  |||||||||||
MPIHNLNHVN MFLQVIASGS ISSAARILRK SHTAVSSAVS NLEIDLCVEL VRRDGYKVEP  60
TEQALRLIPY MRSLLNYQQL IGDIAFNLNK GPRNLRVLLD TAIPPSFCDT VSSVLLDDFN 120
MVSLIRTSPA DSLATIKQDN AEIDIAITID EELKISRFNQ CVLGYTKAFV VAHPQHPLCN 180
ASLHSIASLA NYRQISLGSR SGQHSNLLRP VSDKVLFVEN FDDMLRLVEA GVGWGIAPHY 240
FVEERLRNGT LAVLSELYEP GGIDTKVYCY YNTALESERS FRRFLESARQ RLRELGRQRF 300
             310        320        330        340        350        360
      |||||||||||  |||||||||||  |||||||||||  |||||||||||  |||||||||||  |||||||||||
DDAPAWQPSL GRNCAAPLRP EGAPRPPTRR TRVERSPADA GRGW 344
```

Fig. 7I

Sequence: 34B12 ORF 1 From:: 1 To: 759

```
              10         20         30         40
    ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴

ATGCGGCGAG ATGCCCCATC CGACGCCGGC TTCCACCAGA  40
CGCAGCATGT CGTCGAAGTT TTCCACGAAG AGCACCTTGT  80
CGCTGACCGG CCGCAGCAGG TTCGAATGCT GCCCGGAGCG  120
GCTGCCGAGG CTGATCTGCC GGTAATTGGC CAGGCTCGCG  160
ATGCTGTGCA GGGAGGCATT GCACAACGGG TGCTGCGGAT  200

             210        220        230        240
    ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴

GGGCGACGAC GAACGCCTTG GTGTAGCCGA GCACGCACTG  240
GTTGAAGCGG GAGATCTTCA GTTCCTCGTC GATGGTGATG  280
GCGATATCGA TTTCCGCGTT GTCCTGCTTG ATCGTCGCCA  320
GGCTATCGGC GGGCGAGGTG CGTATCAGGC TGACCATGTT  360
GAAATCGTCG AGCAGTACGC TGCTCACCGT ATCGCAGAAC  400

             410        420        430        440
    ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴

GACGGCGGGA TGGCGGTGTC CAGCAGCACC CGGAGATTGC  440
GCGGACCCTT GTTGAGATTG AAGGCGATGT CGCCGATCAG  480
CTGCTGGTAG TTCAGCAGGC TGCGCATGTA AGGGATCAGG  520
CGAAGCGCCT GCTCGGTGGG TTCGACCTTG TAGCCGTCCC  560
GACGGACCAG CTCCACGCAC AGGTCGATTT CCAGGTTGCT  600

             610        620        630        640
    ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴

GACCGCCGAG CTGACCGCGG TGTGCGACTT GCGCAGGATC  640
CGCGCAGCGG AGGAAATCGA ACCGGAGGCG ATGACCTGGA  680
GGAACATGTT CACGTGATTC AGGTTATGAA TAGGCATCCC  720
TTATTCCTTT TATTGGGTGG CGCGTGCCGC TTCCCTTGA   759
```

Fig. 7J

Sequence: 34B12 ORF 1 S PROTEIN From:: 1 To: 253

```
                10          20          30          40
     ⌐ı ı ı ı l ı ı ı ı ı⌐ ı ı ı ı l ı ı ı ı ı⌐ ı ı ı ı l ı ı ı ı ı⌐ ı ı ı ı l ı ı ı ı ı⌐
MRRDAPSDAG FHQTQHVVEV FHEEHLVADR PQQVRMLPGA 40
AAEADLPVIG QARDAVQGGI AQRVLRMGDD ERLGVAEHAL 80
VEAGDLQFLV DGDGDIDFRV VLLDRRQAIG GRGAYQADHV 120
EIVEQYAAHR IAERRRDGGV QQHPEIARTL VEIEGDVADQ 160
LLVVQQAAHV RDQAKRLLGG FDLVAVPTDQ LHAQVDFQVA 200
                210         220         230         240
     ⌐ı ı ı ı l ı ı ı ı ı⌐ ı ı ı ı l ı ı ı ı ı⌐ ı ı ı ı l ı ı ı ı ı⌐ ı ı ı ı l ı ı ı ı ı⌐
DRRADRGVRL AQDPRSGGNR TGGDDLEEHV HVIQVMNRHP 240
LFLLLGGACR FP. 253
```

Fig. 7K

pho34B12 ORF1 (L-S) SEQ ID NO:107

```
  1 MARARPCVSI SVTRCSTASL TCASRATTRW WSVACRKVST AWPMPMARSW
 51 RWPIRAIDSW ACNSIPSRFS PPTASVCWRT LYSGAARWRS RSAFGPERRC
101 AVSTEARLPG RRIVETLAAQ FAQALAGAFE KATEAALGFQ RGVVVAIHLG
151 VDAARFVQFA EDCQGTVAQA FLDEIMRRDA PSDAGFHQTQ HVVEVFHEEY
201 LVADRPQQVR MLPGAATEAD LPVIGQARDA VQGGIAQRVL RMGDDERLGV
251 AEHALVEAGD LQFLVDGDGD IDFRVVLLDR RQAIGGRGAY QADHVEIVEQ
301 YAAHRIAERR RDGGVQQHPE IARTLVEIEG DVADQLLVVQ QAAHVRDQAK
351 RLLGGFDLVA VPTDQLHAQV DFQVADRRAD RGVRLAQDPR SGGNRTGGDD
401 LEEHVHVIQV MNRHPLFLLL GGACRFP*
```

Fig. 8

phoB12 ORF2 SEQ ID NO: 108

```
  1 MPIHNLNHVN MFLQVIASGS ISSAARILRK SHTAVSSAVS NLEIDLCVEL
 51 VRRDGYKVEP TEQALRLIPY MRSLLNYQQL IGDIAFNLNK GPRNLRVLLD
101 TAIPPSFCDT VSSVLLDDFN MVSLIRTSPA DSLATIKQDN AEIDIAITID
151 EELKISRFNQ CVLGYTKAFV VAHPQHPLCN ASLHSIASLA NYRQISLGSR
201 SGQHSNLLRP VSDKVLFVEN FDDMLRLVEA GVGWGIAPHY FVEERLRNGT
251 LAVLSELYEP GGIDTKVYCY YNTALESERS FRRFLESARQ RLRELGRQRF
301 DDAPAWQPSL GRNCAAPLRP EGAPRPPTRR TRVERSPADA GRGW*
```

Fig. 9

36A4 SEQ ID NO: 109

```
  1 AAGGGTTTTG GCGGGGTCAT CCGAGTGACC CTGAGCATGC TCCTGGCGAT
 51 CTTCTTGTCG GTGCTGCTGG CGCCGGTGCG CATGCTGTTC CACACCCGCT
101 TCGTGCTGGC CGCCTTCCTC GGCTGGTC
```

Fig. 10

36A4 SEQ ID NO:110

```
  1 KGFGGVIRVT LSMLLAIFLS VLLAPVRMLF HTRFVLAAFL GW
```

Fig. 11

contig 2507 SEQ ID NO: 111

```
   1 CTACTGGGGC AAGCTGAAGA CGCCGTTCAA GCTGAGCTTC TATCACCAGG
  51 GCATGCACTT CGACACGCCG GTGAAGATCA ACGAGGTGAC CGCTACCACG
 101 GTCAAGCCGA TCAAGTACGA TCGCACCAAG TTCGATTTCG GATCCCTGAA
 151 GTTCGACGAG AATGCCACCA AGGATCTCGG CTATGCCGGT TTCCGCGTGC
 201 TCTATCCGAT CAACAAGGCC GACAAGCAGG ACGAGATCGC CACCTTCCTT
 251 GGCGCGAGCT ACTTCCGCGT GGTCGGCAAG GGCCAGGTCT ACGGTCTGTC
 301 GGCGCGCGGC CTGGCGATCG ATACCGCGCT GCCTTCGGGC GAAGAGTTCC
 351 CGCGCTTCCG CGAATTCTGG ATCGAGCGGC CGAAGCGCAG GACAAGCAAC
 401 TGGTGATCTA CGCCCTGCTC GACTCGCCGC GGGCCACCGG CGCCTACCGC
 451 TTCGTGCTGC GTCCGGGCAA GGATGCGGTG ATGGATGTCC AGGCCCGCGT
 501 GTTCCTCCGC GACAAGGTCA GCAAGCTGGG CCTGGCGCCG CTGACCAGCA
 551 TGTACCTGTT CGGCTCCAAC CAGCCGTCCG AGCAGCACAA CTTCCGGCCC
 601 GAGCTGCATG ACTCCAGCGG CCTGCAGATC CATGCCGGCA ACGGCGAGTG
 651 GCTGTGGCGT CCGCTGAACA ATCCGAAGCA CCTGTCGGTG AGCACCTTCA
 701 GCGTGGAGAA CCCGAAAGGC TTCGGCCTGC TCCAGCGCGG CCGCGAGTTC
 751 TCCCGCTACG AAGACCTGGA TGACCGCTAC GACCTGCGTC CGAGTGCCTG
 801 GATCGAGCCG AAGGGCGACT GGGGCAAGGG CACCGTGGAA CTGGTGGAAA
 851 TCCCGACCCC GGACGAAACC AACGACAATA TCGTCGCGTT CTGGAACCCC
 901 GAGACCCAGC CTGAGGTCGG AAAGCCGCTG GACTTCGCCT ACCGCCTGCA
 951 CTGGACCATG GATGAAGACG AGCTGCACGA CCCGAAATCC TCCTGGGTCA
1001 AGCAGACCAT GCGCTCGGTC GGCGACGTGA AGCAGAAGAA CCTGATCCGC
1051 CAGCAGGACG GCAGCACCGC CCTGGTCGTC GACTTCGAAG GGCCGGCCCT
1101 GAAGGACCTG GCGCCGGACG CGCCGGTGAC CACCCAGGTC AGCACCGACA
1151 GCAACGCCGA GGTGGTGGAG AACAGCCTGC GTTACAACCC GGTCCTGAAA
1201 GGCTGGCGCC TGACGCTGCG GATCAAGGTC AAGGATCCGA AGAAGCCGGT
1251 GGAAATGCGC GCGGCGCTGG TCGACGAGGC GCAGAAGCCA CTGAGCGAAA
1301 CCTGGAGCTA TCAGCTGCCT GCCGATGAAT AACCCATCCA CTACGAAAGC
1351 ACCGCTGGCC GACTACCTCG CTCATCTTCC CCTGGCGGAA GAGGAGCGGG
1401 AGCGCCTTGG CGAGTCCGCT TCCTTCTCCG AGCTGCACGC TCGCCTGGCG
1451 GGAGCGGAAG GCGCCGCTGC CGATGCCGGG GGCGATCCCG CCCTGGCCTC
1501 GGTACGCGCC CGCCTGCAGC TGGGCACCCC TGAGCTGGAC GACGCCGAGA
1551 TGTTCGGCGT CGACGCCCAG GGTCGCACCT TCCTCAAGAT TTCCCCGCCG
1601 ATCCGCCGTA CCAAGGTGAT TCCCGAGCCC TGGCGCACCA ACATCCTGGT
1651 GCGCGGCTGG CGTCGGCTGA CCGGACGCAG CAACCCGCCC AAGCCCAAGC
1701 GTGCCCTGCC GCGGGCCCGC TGGCAGCGGG TCGGCTCGCT GCGCCGGTTC
1751 ATCCTGCTGT TGTTGATGCT GGCGCAGACC TCGGTCGCCA CCTACTACAT
1801 GAAAGGCATC CTGCCCTACC AGGGCTGGGC CTTCGTCGAC CTGGAGGAGC
1851 TGGCCCAGCA GAGCCTGCTG GATACCGTCC AGCAGGTGCT GCCCTATGTC
1901 ATCCAGTTCG GCATCCTGGC GCTCTTCGCG ATCCTCTTCT GCTGGGTCTC
1951 GGCCGGCTTC TGGACCGCGC TGATGGGCTT CTGGGAGCTG CTCACCGGGC
2001 GTGACCGCTA CCGGATCTCC GGCAGCAGCG CCGGCAGCGA GCCGATCGCC
2051 GCCGACGCCC GCACGGCGAT CGTCATGCCG ATCTGCAACG AAGACGTGCC
2101 GCGGGTATTC GCCGGCCTGC GGGCGACCGT CGAGTCGATG GCCGCCACCG
2151 GCGAGATGGA GCGCTTCGAC TTCTTCGTCC TCAGCGACAC CAACGACCCG
2201 GATATCGCCG TCGCCGAGCA GCAGGCCTGG CTCGAGCTGT GCCGCGAGAC
2251 CAAGGGCTTC GGCAAGATCT TCTACCGTCG CCGCCGGCGC CGGGTGAAGC
2301 GCAAGAGCGG CAACATCGAC GACTTCTGCC GGCGCTGGGG CGGCGACTAC
2351 CGCTACATGG TGGTGATGGA CGCCGACAGC GTGATGAGCG GCGACTGCCT
```

Fig. 12A

```
2401 GGCCAAGCTG GTACGCCTGA TGGAGGCCAA TCCTGAGGCG GGGATCATCC
2451 AGACCGCGCC GAAGGCTCCG GCATGGACAC CCTGTATGCG CGCATGCAGC
2501 AGTTCGCCAC CCGCGTCTAC GGCCCGCTGT TCACCGCCGG CCTGCACTTC
2551 TGGCAACTCG GCGAGTCGCA CTACTGGGGC CACAACGCGA TCATCCGCAT
2601 GCAGCCCTTC ATCGACCACT GCGCCCTGGC GCCGTTGCCG GGCAAGGGCT
2651 CGTTCGCCGG CGCGATCCTG TCCCACGACT TCGTCGAGGC TGCGTTGATG
2701 CGCCTTGCCG GCTGGGGCGT GTGGATCGCC TACGACTTCG ACGGCAGCTA
2751 CGAAGAACTG CCGCCGAACC TGCTCGACGA ACTCAAGCGC GACCGCCGCT
2801 GGTGCCACGG CAACCTGATG AACTTCCGCC TGTTCCTGGT CAAGGGCATG
2851 CACCCGGTGC ACCGCGCGGT GTTCCTCACC GGGGTCATGT CCTACCTGTC
2901 GGCGCCGTTG TGGTTCTTCT TCCTGGTGCT TTCCACCGCG CTGCTGGCGG
2951 TGCACCAACT GATGGAGCCG CAGTACTTCC TGGAACCGCG GCAGCTGTTC
3001 CCGATCTGGC CGCAGTGGCA TCCGGAGAAG GCCATCGCGT TGTTCTCCAC
3051 CACCCTGACC CTGTTGTTCC TGCCCAAGCT GCTCAGCGTA ATGCTGATCT
3101 GGGCCAAGGG CGCCAAGGGT TTCGGCGGGG TGATCCGGGT GACCCTGAGC
3151 ATGCTCCTGG AGATGTTCTT CTCGGTGCTG CTGGCGCCGG TGCGCATGCT
3201 CTTCCACACC CGCTTCGTGC TGGCCGCCTT CCTCGGCTGG TCGGTGCAGT
3251 GGAACTCGCC GCAGCGCGAC GACGACGCCA CGCCCTGGAG CGAGGCGATC
3301 CGCCGGCACG CAATGCAGAC CCTGCTGGGT ATCGCCTGGA CCCTGCTGGT
3351 GGCCTGGCTC AACCCGCGCT TCCTGTGGTG GCTGTCGCCG ATCGTCGGTT
3401 CGCTGATCCT GTCGATCCCG GTATCGGTGA TCTCCAGCCG GGTGAAGCTG
3451 GGCCTGCGGG CCCGCTACGA AAAGCTGGTC CTGATCCCGG AGAGTACGAC
3501 ACGCCGCGCG ACTGCGCGCC ACCGACGAGT ACACCTACGA GAACCGCTGG
3551 CATGCGCTCA AGGATGGCTT CCTCAAGGCC GCCGTCGATC CGTTGCTCAA
3601 CGCCCTGGCC TGCGCCATGG GCACGGCTCG CCACAACCGT GCGCAGGCCA
3651 TCGAGACGGT GCGTGGCGAG CGTATCGGCA AGGCCATCGA TAAGGGCCCG
3701 GAACAGCTCG ACGGCGCCAC GCGCCTGGCT CTGTTGAGTG ACCCGGTAGC
3751 ACTTTCGCGC CTGCATACGC GGGTCTGGGA AGAGGACCGC GACGACTGGC
3801 TCGGCCGCTG GCGCAAGGCC GAGGCGGACG ACCCCCACGC CGCCAGCGTA
3851 CCGCTGGCCC AGGTAGTGCC CGGCGACGCC GGCCTGCTGC CCGCCGCCCA
3901 GTCCTGATCC CATGCCCCCG GCGGAACGCC GCCGGGGGCA TGGGTCTGTT
4001 TCTTGCCTGT TTTCCCCGTG CGGCGCTGCT GTTACCCTGC GCCGGCAATC
4051 CAGAAAGTCT CGTATCGTTC GCCAGCTGAG GTACTATCGG CCGCCTTTTG
4101 CGCAGCCGGT CATGGCCTGC TGCCCGCCCG GGACGGCGAC ACGACGAGAG
4151 CATCCGTTCG ACGACTGTGT TTCTAAGACT GCTGGGGATT GGGGAATGAA
4201 AAAGTATCTT GCTTCATTGG TTCTGGGCGT CTGCGCCCTG GTGGGCGTGG
4251 CTTCGGTCCA GGCGGCCGGC GCGGTGGAGG ACGCGGTCAA GCGCGGCACC
4301 CTGCGGGTCG GCATGGACCC GACCTACATG CCGTTCGAGA TGACCAACAA
4351 GCGTGGCCAG ATCATCGGCT TCGAAGTCGA CCTGCTCAAG GCCATGGCCA
4401 AGTCCATGGG CGTCAAGCTG GAGCTGGTCT CCACCAGCTA CGACGGCATC
4451 ATCCCGGCGC TGCTGACCGA CAAGTTCGAC ATGATCGGCT CGGGCATGAC
4501 CCTGACCCAG GAGCGCAACC TGCGCCTGAA CTTCTCCGAG CCCTTCATCG
4551 TGGTCGGCCA GACCCTGCTG GTGCGCAAGG AACTGGAAGG CAAGATCAAG
4601 TCCTACAAGG ACCTGAACGA TCCGCAGTAC AGCATCACCT CGAAGATCGG
4651 CACCACCGGT GAGATCGTTG CCCGCAAGCT GATCAGCAAG GCCAAGTACC
4701 ACGGCTTCGA CAACGAGCCG GAAGCGGTGA TGGACGTGGT CAACGGCAAG
4751 GCCGACGCCT TCATCTACGA CTCGCCCTAC AACGTGGTGG CGGTGAGCAA
4801 GTTCGGCGCC GGCAAGCTGG TCTACCTCGA CCAGCCGTTC ACCTACGAGC
4851 CGCTGGCGTT CGGCCTGAAG AAAGGCGACT ACGACAGCAT CAATTTCATC
4901 AACAACTTCC TCCATCAGAT CCGCGAAGAC GGCACCTATC AGCGCATCCA
```

Fig. 12B

```
4951 CGACAAGTGG TTCAAGAACA CCGAGTGGCT GAAGGAAATG GAATGAACCG
5001 CTGACGGCCC CCGCGAAGGG GGCCGTCGTA CCTGCGCATT CCATCGTTCG
5051 AGAGAGTTTC CGTGACCAAG AAGAAACGTT CCGTCTGGCC CTGGCACCTG
5101 CTGACCGGGC TGATCCTGCT GGTCATGGCC TGGGCGCTGT GGTTCTCCAC
5151 CTCGCTGATT TCCTATGAAA TGGCGTGGGA CCGCGTTTCC GAGTACTTCG
5201 CTACCAGGCC GAGGAGCCGT TACGGGCCAA CGAGATCGGC CGGGTCGAGG
5251 CTATCGAGGA ACAGGGCAGG GACGCGCGCG TCACGCTGCT TGGCGAGACG
5301 GCGAGAAGCA GGTCGTGACC GTTGCCCAGG ACAGCCTGCA ATTCTCCGAA
5351 GCGACGACGT GGCCGAGGGC GACGCGGTCG GGGTGACCCG CCACTGGGCC
5401 GCCGGCACTG CTCTGGGGCC TGTGGACCAC CCTCTGGCTA TCGCTGGTGT
5451 CCGGTGCCAT CGGTCTGGCT ATCGGCCTGG TCGCCGGCCT CTGCCGGCTG
5501 TCGAAGAACC CGACCCTGCA CGACCTGTCG ACGATCTACG TCGAGCTGGT
5551 GCGCGGCACG CCGTTGCTGG TGCAGATCTT CATCTTCTAC TTCTTCATCG
5601 GCACCGTGCT CAACCTGTCC CGCGAGTTCG CCGGGGTTGC GGCGCTGGCG
5651 CTGTTCACCG GCGCCTACGT GGCCGAGATC ATCCGGGCCG GCGTGCAGTC
5701 CATCGCCCGC GGACAGAACG AGGCCGCCCG CTCCCTGGGC CTGAACGCCG
5751 GCCAGTCGAT GCGCTACGTG ATCCTGCCGC AGGCTTCAAG CGCGTGCTGC
5801 CGCCGCTGGC CGGGCAGTTC ATCAGCCTGG TCAAGGACAC CTCGCTGGTC
5851 TCGGTGATCG CCATCACCGA ACTGACCAAG AGCGGCCGCG AGGCGATCAC
5901 CCACTTCGTT CTCCAACTTC GAGATCTGGT TTCTGCGTCG CCGCGTTGTA
5951 CCTGCTGTTG AACCTGCCCC TTTCGCACAT GGCATCCCGA CTGGAGCGGA
6001 GGCTCGGACA AAGTGATTGA AGTACGCAAC CTGCTGAAGG TCTTCGATAC
6051 CCGCGGCCAG GTAGTGCGCG CGGTGGACGA CGTGAGTACC CGCGTGGCCA
6101 GGGGCGAGGT ACTGGTGGTG ATCGGTCCGT CCGGTTCCGG CAAGTCGACC
6151 TTCCTGCGCT GCCTGAACGG CCTGGAGGAG TTCGACGAAG GCTCGGTGAG
6201 CATCGACGGC GTCGACCTGG CCGACCCGAG GACCGACATC AATGCCTACC
6251 GCCGCGAAGT CGGCATGGTG TTCCAGCATT TCAACCTGTT CCCGCACATG
6301 ACCGTGCTCG AGAACCTCTG CCTGGCCCAA CGCGTGGTGC GCAAGCGCGG
6351 CAAGGCCGAG CGCGAGGCCA AGGCGCGGGC GCTGCTGGCC AAGGTCGGCA
6401 TCGGGCAGAA GGCCGACGAA TATCCCTCGC GCCTGTCCGG CGGCCAGCAG
6451 CAGCGCGTGG CGATCGCTCG CGCGTTGTGC ATGGACCCCA AGGTGATGCT
6501 GTTCGACGAA CCGACCTCGG CGCTCGATCC GGAGATGGTC GGCGAAGTCC
6551 TCGACGTCAT GAAGACCCTG GCCGTGGAAG GCATGACCAT GGTCTGCGTG
6601 ACCCACGAGA TGGGCTTTGC CCGCGAAGTG GCCGACCGCG TGCTGTTCTT
6651 CGACCACGGC AAGCTGCTGG AGGACGCGCC GCCGGCGCAG TTCTTCGACA
6701 ATCCGCAGGA CCCGCGGGCC CAGGCCTTCC TCCGCCAGGT CCTCTAGTAC
6751 CGCGCTAGGC GAACGGCTTG CCCGGCGGCG GCAGGAGCGA CGTCGGACTC
6801 TGCCGCGCGG CCGGCTGGAT ATCGTTGTCC TCCAGCCAGT CCAGCGCCCA
6851 TTCGCGCAGG CGCTCGTTCT GGTAGCGGTA CCAGTCCTGC AACAGTTCCG
6901 GGTACTCCAT CAGAGAGTGC TTGAAGGCCT TGAACGGCTT GCGGCTCTGC
6951 AGCGCGTTG
```

Fig. 12C

23A2 DNA SEQ ID NO:112

```
  1 CGAGGTTTCC GTCTACGAAG GCACCGGCTC GGTCACCATC CGCGCCGTGT
 51 TCCCCAACCC GAACAACGAG CTGCTCCCCG GCATGTTCGT TCACGCGCAG
101 TTGCAGG
```

Fig. 13

23A2 peptide SEQ ID NO:113

```
  1  EVSVYEGTGS VTIRAVFPNP NNELLPGMFV HAQLQ
```

Fig. 14A

SEQ ID NO:148
  DNA flanking the 23A2 locus.
  mexA partial sequence, mexB partial sequence

```
   1 ggccaggcaa acgcgatggc caccgtgcaa cagctcgacc cgatctacgt cgacgtcacc
  61 cagccgtcca ccgccctgtt gcgcatgcgc cgcgaactgg ccagcggcca gttggagcgc
 121 gccggcgaca acgctgcgaa ggtctccctg aagctggagg acggtagcca atacccgctg
 181 gaaggccgcc tcgaattctc cgaggtttcc gtcgacgaag gcaccggctc ggtcaccatc
 241 cgcgccgtgt tccccaaccc gaacaacgag ctgctgcccg gcatgttcgt tcacgcgcag
 301 ttgcaggaag gcgtcaagca gaaggccatc ctcgctccgc agcaaggcgt gacccgcgac
 361 ctcaagggcc aggctaccgc gctggtggtg aacgcgcaga acaaggtcga gctgcgggtg
 421 atcaaggccg accgggtgat cggcgacaag tggctggtca ccgaaggcct gaacgccggc
 481 gacaagatca ttaccgaagg cctgcagttc gtgcagccgg gtgtcgaggt gaagaccgtg
 541 ccggcgaaga atgtcgcgtc cgcgcagaag gccgacgccg ctccggcgaa aaccgacagc
 601 aagggctgat caagggggatt cgtaatgtcg aagtttttca ttgataggcc cattttcgcg
 661 tgggtgatcg ccttggtgat catgctcgcg ggcggcctgt cgatcctcaa tctgccggtc
 721 aaccagtacc cggccatcgc cccgccggcc atcgccgtgc aggtgagcta cccgggcgcc
 781 tcggccgaga cggtgcagga caccgtggtc caggtgatcg agcagcagat gaacgggatc
 841 gacaatctgc gctacatctc ctcggagagt aactccgacg gcagcatgac catcaccgtg
 901 accttcgaac agggcaccga ccccgacatc gcccaggtcc aggtgcagaa caagctgcaa
 961 ctggccaccc cgctgctgcc gcaggaagtg cagcgccagg ggatccgg
```

Fig. 14B

SEQ ID NO:149
  PA14 mexA

G QANAMATVQ QLDPIYVDVT QPSTALLRMR RELASGQLER AGDNAAKVSL KLEDGSQYP   LEGRLEFSE
VSVDEGTGS VTIRAVFPN PNNELLPGM FVHAQLQEG VKQKAILAP QQGVTRDLK GQATALVVN
AQNKVELRV IKADRVIGD KWLVTEGLN AGDKIITEG LQFVQPGVE VKTVPAKNV ASAQKADAA PAKTDSKG

Fig. 14C

SEQ ID NO:150
  PA14 mexB

MSKFFIDRPIFAWVIALVIMLAGGLSILNLPVNQYPAIAPPAIA
VQVSYPGASAETVQDTVVQVIEQQMNGIDNLRYISSESNSDGSMTITVTFEQGTDPDI
AQVQVQNKLQLATPLLPQEVQRQGIR

Fig. 14D

PAO1 Phenazine operon SEQ ID NO:114

```
   1 GCAAGCTCAA CTCCAGCAAC AAGGCGGAGG CCACCATGAA GGCTTACGCC
  51 ATCGGCCTGC TCAACTGAAT CGACGCCTCG TCGCCTAGCG AGGCCGCCGC
 101 GCAAGCGTCC GGCCATTCAC CGAATGGCCG GATAGCGTTT GCGCCGGTCG
 151 CCTGAGCGCA CGCTTCCCAC CGGCAGCGTT TCCCCGCTGC CCCCTTCGCC
 201 ATTGCGCCCG TCCTCATGTT GTCCGGACGC TAGTCGAACT TTCCGGGCGC
 251 CTGGCAAACC GGCCAAAGAA TAGAACGGAA TCGATGCCCA CACCTTTAAT
 301 TTTTAAGGGT TTTTCCTTTT CAAAAACCGT TATTAAGTTT TCCCCTTTAA
 351 ATCTTGGTAC AACTGGGTTC AGGCGAAACT TCGGTCATGC CATTCGGCAT
 401 TAGTTAAACT TTGAGACTCT CCAAGCGGGA ATTTTTGCCG GAACAGCTTC
 451 ACGGCATTTC TCCGCTTTCA TCCCGATGTT TCTTTCCGTT ATGATTCCAG
 501 TCGATTCGAA CTGCCGGAGT TCCCACCATT CGAGATTACC AACGTTGAAA
 551 AGGGTTTACC GACAACCTGG AATTGCGTCG GCGCAACCGT GCCACGGTCG
 601 AGCACTACAT GCGCATGAAG GGGGCCGAAC GGTTACAGCG GCACAGCCTG
 651 TTCGTCGAGG AGGCTGCGCC GGCAACTGGA CCACGGAAAG CGGCGAACCC
 701 CTGGTTTTCC GGGGCCATGA GAGCCTCAGG CGGCTCGCCG AGTGGCTCGA
 751 GCGCTGCTTC CCCGACTGGG AGTGGCACAA CGTGCGGATC TTCGAGACCG
 801 AGGATCCGAA CCACTTCTGG GTCGAGTGCG ACGGGCGCGG CAAGGCGCTG
 851 GTCCCGGGGT ATCCGCAGGG CTATTGCGAG AACCACTACA TCCATTCCTT
 901 CGAACTCGAG AACGGCCGGA TAAAACGCAA TCGCGAGTTC ATGAACCCGA
1001 TGCAGAAATT GCGTGCATTG GGAATAGCCG TTCCACAAAT AAAACGTGAC
1051 GGTATTCCCA CCTGATTAAT GTCTATTCCA ATTCAAGAGG AGATATGACG
1101 ATGCTCGATA ATGCCATTCC TCAAGGTTTC GAAGACGCCG TGGAGTTGCG
1151 CAGGAAGAAT CGCGAGACGG TGGTCAAGTA TATGAACACC AAAGGCCAGG
1201 ATCGCCTGCG CCGCCATGAA CTTTTCGTCG AGGACGGCTG TGGCGGTTTA
1251 TGGACCACCG ATACCGGCTC GCCCATCGTC ATTCGTGGCA AGGACAAGCT
1301 GGCCGAGCAC GCGGTGTGGT CGCTGAAATG CTTCCCGGAT TGGGAGTGGT
1351 ACAACATCAA GGTCTTCGAG ACCGACGATC CCAACCACTT CTGGGTCGAG
1401 TGCGACGGCC ACGGCAAGAT CCTCTTCCCC GGCTATCCCG AGGGCTACTA
1451 CGAGAACCAC TTCCTGCATT CCTTCGAGCT GGACGACGGC AAGATCAAGC
1501 GCAACCGCGA ATTCATGAAC GTCTTCCAGC AATTGCGCGC CCTGAGCATT
1551 CCGGTCCCGC AGATCAAACG CGAAGGCATT CCCACCTGAG GCCATCCTGG
1601 AAGGGGTGAA CTATGGACGA TCTATTGCAA CGCGTACGGC GCTGCGAAGC
1651 GCTGCAGCAA CCCGAATGGG GCGATCCGTC GCGCCTGCGC GACGTGCAGG
1701 CGTACCTGCG CGGCAGTCCG GCGCTGATCC GCGCCGGCGA CATCCTGGCC
1751 CTGCGCGCGA CCCTGGCGCG GGTCGCCCGC GGCGAGGCGC TGGTGGTGCA
1801 GTGCGGCGAC TGCGCCGAGG ACATGGACGA CCACCATGCC GAGAACGTGG
1851 CGCGCAAGGC CGCCGTGCTG GAACTGCTGG CCGGCGCCCT GCGCCTGGCC
1901 GGCCGGCGGC CGATAGATCC GCGTCGGGCG CATCGCCGGG CAGTACGCCA
1951 AGCCGCGTTC CAAGCCGCAC GAGCAGGTCG GCGAGCAGAC CCTGCCGGTC
2001 TATCGCGGCG ACATGGTCAA CGGCCGCGAG GCCCATGCCG AACAGCGCCG
2051 GGCCGATCCG CAGCGGATCC TCAAGGGCTA TGCGGCGGCG CGCAACATCA
2101 T
```

<p style="text-align: center;">Fig. 15</p>

3E8 sequence SEQ ID NO:115

```
   1 CGGCGCCGAG GATCCGCTGT TCGAGTTAGG CGCAAGCGTC CGGCCATTCA
  51 CGGAATGGCC GGATAGCGTT TGCGCCGGTT GCTTGAGCGC AGCTTCCCAC
 101 CGGCAGGGTT TCCCCGCTGC CCCTTTCGCC ATTGCGCCGT CCTCTTGTTG
 151 TCCGGCACGC TAGTGCAACT TTCCGGACGC TTGGCAAACC GGCCAAAGAA
 201 TAGAACGGAA TCGATGCCCC ACACCTGTAA TTTTTAAGGG GTTATGGCTA
 251 TTGCAAAAAA GCGTTTATAA GTTTGTCCCC TGTCAAATCT GGTTACAACT
 301 GGGTTTCAGG CGAAACATTC GGTCATGGCA ATTCGGCATT AGTTGAAACT
 351 TTGGAGACGC TCCGAAGCGG GCAACTTTTG CCCGGAAAAA GTTTCACGGC
 401 AATTTTTCCG GCCTGTCATC CCGATGTCTT CTTTCCAGTA TGGATGCCAG
 451 TCGATTCGAA CTGGCGGAGA TTCGCACCAT GCGAGAGTAC CAACGGTTGA
 501 AAGGGTTTAC CGACAACCTG GAATTGCGGC GGCGCAACCG TGCCACGGTC
 551 GAGCACTACA TGCGCATGAA GGGGGCCGAA CGGTTGCAGC GGCACAGCCT
 601 GTTCGTCGAG GACGGCTGCG CCGGCAACTG GACCACGGAA AGCGGCGAAC
 651 CCCTGGTTTT CCGGGGCCAT GAGAGCCTCA GGCGGCTCGC CGAGTGGCTC
 701 GAGCGCTGCT TCCCCGACTG GGAGTGGCAC AACGTGCGGA TCTTCGAGAC
 751 CGAGGATCCG AACCACCTCT GGGTCGAGTG CGACGGGCGC GGCAAGGCGC
 801 TGGTCCCGGG GTATCCGCAG GGCTATTGCG AGAACCACTA CATCCATTCC
 851 TTCGAACTCG AGAACGGCCG GATAAAACGC AATCGCGAGT TCACGAACCC
 901 GATGCAGAAA TTGCGTGCAT TGGGAATAGC CGTTCCGCAA ATAAaACGTG
 951 ACGGCATTCC CACCTGATTA ATGATTATTC CAATTCAAGA GGAGATATGA
1001 CGATGCTCGA TAATGCTATT CCCCAAGGTT TCGAAGACGC CGTGGAGTTG
1051 CGCAGGAAGA ATCGCGAGAC GGTGGTCAAG TATATGAACA CCAAAGGCCA
1101 GGATCGCCTG CGCCGCCATG AACTTTTCGT CGAGGACGGC TGTGGCGGTT
1151 TATGGACCAC CGATACCGGC TCGCCCATCG TCATTCGTGG CAAGGACAAG
1201 CTGGCCGAGC ACGCGGTGTG GTCGCTGAAA TGCCTTCCCG GATTGGGAGT
1251 GGTACAACAT CAAGGT
```

Fig. 16A

3E8 SEQUENCE TAG SEQ ID NO:160

```
   1 tatggatgcc agtcgattcg aactggcgga gattcgcacc atgcgagagt accaacggtt
  61 gaaagggttt accgacaacc tggaattgcg gcggcgcaac cgtgccacgg tcgagcacta
 121 catgcgcatg aaggggggccg aacggttgca gcggcacagc ctgttcgtcg aggacggctg
 181 cgccggcaac tggaccacgg aaagcggcga acccctggtt ttccggggcc atgagagcct
 241 caggcggctc gccgagtggc tcgagcgctg cttccccgac tgggagtggc acaacgtgcg
 301 gatcttcgag accgaggatc cgaaccacct ctgggtcgag tgcgacgggc gcggcaaggc
 361 gctggtcccg gggtatccgc agggctattg cgagaaccac tacatccatt ccttcgaact
 421 cgagaacggc cggataaaac gcaatcgcga gttcacgaac ccgatgcaga aattgcgtgc
 481 attgggaata gccgttccgc aaataaaacg tgacggcatt cccacctgat taatgattat
 541 tccaattcaa gaggagatat gacgatgctc gataatgcta ttccccaagg tttcgaagac
 601 gccgtggagt tgcgcaggaa gaatcgcgag acggtggtca agtatatgaa caccaaaggc
 661 caggatcgcc tgcgccgcca tgaacttttc gtcgaggacg gctgtggcgg tttatggacc
 721 accgataccg gctcgcccat cgtcattcgt ggcaaggaca agctggccga gcacgcggtg
 781 tggtcgctga aatgcttccc ggattgggag tggtacaaca tcaaggtctt cgagaccgac
 841 gatcccaacc acttctgggt cgagtgcgac ggccacggca agatcctctt ccccgggtat
 901 cccgagggtt actacgagaa cccacttcctg cattccttcg agctggacga cggcaagatc
 961 aagcgcaacc gcgaattcat gaacgtcttc cagcaattgc gcgccctgag cattccggtc
1021 ccgcagatca aacgcgaagg cattcccacc tgaggccatc ctggaagggg tgaactatgg
1081 acgatctatt gcaacgcgta cggcgctgcg aagcgctgca gcaacccgaa tggggcgatc
1141 cgtcgcgcct gcgcgacgtg caggcgtacc tgcgcggcag tccggcgctg atccgcgccg
1201 gcgacatcct ggccctgcgc gcgaccctgg ccgggtcgcc cgcggcgagg cgctggtggt
1261 gcagtgcggc gactgcgccg aggacatgga cgaccacca
```

Fig. 16B

3E8 phzA   SEQ ID NO:116

```
  1 MREYQRLKGF TDNLELRRRG SAVRVRRKRP AIHGMAGZRL RRLLERSFPP
 51 AGFPRCPFRH CAVLLLSGTL VQLSGRLANR PKNRTESMPH TCNFZGVMAI
101 AKKRLZVCPL SNLVTTGFQA KHSVMAIRHZ LKLWRRSEAG NFCPEKVSRQ
151 FFRPVIPMSS FQYGCQSIRT GGDSHHARVP TVERFTDNLE LRRRNRATVE
201 HYMRMKGAER LQRHSLFVED GCAGNWTTES GEPLVFRGHE SLRRLAEWLE
251 RCFPDWEWHN VRIFETEDPN HLWVECDGRG KALVPGYPQG YCENHYIHSF
301 ELENGRIKRN REFTNPMQKL RALGIAVPQI KRDGIPTZLM IIPIQEEIZR
351 CSIMLFPKVS KTPWSCAGRI ARRWSSIZTP KARIACAAMN FSSRTAVAVY
401 GPPIPARPSS FVARTSWPST RCGRZNAFPD WEWYNIK
```

Fig. 17

3E8 phzB SEQ ID NO:117

```
 1 MLDNAIPQGF EDAVELRRKN RETVVKYMNT KGQDRLRRHE LFVEDGCGGL
51 WTTDTGSPIV IRGKDKLAEH AVWSLKCLPG LGVVQHQG
```

Fig. 18A

3E8 PHZA SEQ ID NO:161

```
MREYQRLKGFTDNLELRRRNRATVEHYMRMKGAERLQRHSLFVE
DGCAGNWTTESGEPLVFRGHESLRRLAEWLERCFPDWEWHNVRIFETEDPNHLWVECD
GRGKALVPGYPQGYCENHYIHSFELENGRIKRNREFTNPMQKLRALGIAVPQIKRDGIPT
```

Fig. 18B

PhzB SEQ ID NO:162

```
MLDNAIPQGFEDAVELRRKNRETVVKYMNTKGQDRLRRHELFVEDGCGGLWTTDTGSPIVIRGKDKLAEHAVWSLKCF
PDWEWYNIKVFETDDPNHFWVECDGHGKILFPGYPEGYYENHFLHSFELDDGKIKRNREFMNVFQQLRALSIPVPQIK
REGIPT
```

Fig. 18C

PhzC SEQ ID NO:163

```
MDDLLQRVRRCEALQQPEWGDPSRLRDVQAYLRGSPALIRAGDILALRATLAGSPAARRWWCSAATAPRTWTTT
```

Fig. 18D

PA14 phzR SEQ ID NO:164

phzR DNA sequence : 1161 bp

CGTCGACGAGGCCCGC CATGGGCCAAGGTTTGTTGT CGGGAGGCgCTCCCGACGACGATG
GAGCGTGCGAGAAGAACAATGAGAAAGACCGCCGTGAGGCCCATCGGAGAGCCGTTCTAC
GGTTTCCGCAAAGATCCGGGGCGCCGTCCCCTCCAGCa CAGCGCAGTTCCTGCGCGGCGC
CTCGTGTCCGTGCTCATCGAGAAGTTCTCTTCAGCCTCGTTTCGTCGTCGCCCGGCGGGC
GGCGAATGGGCTCGACCTCGTCCGGAACACCCGCACAGGGCCGGTGGCGATATGTACTTC
CAGGTCCGGCTTGATAAAGGGAATTGTCATGAGTGGATAAGACGGAAACAAAAAAGAATA
AAAACGCTGAAGAACCGAATCCTGCCGGGATCGATTGTTGACTGGTGAAGCTGGCATGCA
TGATGAGAGAGAGGGATATCTCGAGATTTTGTCAAGAATAACAACCGAGGAAGAGTTCTT
CTCCCTGGTTCTCGAGATATGCGGTAATTATGGATTCGAATTCTTTTCATTCGGTGCGCG
GGCGCCTTTCCCGCTGACCGCGCCTAAATATCATTTCCTGTCCAATTACCCAGGGGAATG
GAAAAGCAGATATATCTCCGAAGACTACACATCCATCGACCCGATCGTGCGCCATGGTCT
CCTGGAATACACCCCGCTGATCTGGAATGGCGAAGACTTCCAGGAGAACCGTTTCTTCTG
GGAGGAAGCGCTGCATCACGGCATCCGTCACGGCTGGTCGATCCCGGTCCGCGGCAAGTA
CGGGCTGATCAGCATGCTGTCCCTGGTGCGTTCCAGCGAGAGCATCGCCGCTACGGAAAT
CCTGGAGAAGGAATCCTTCCTGCTCTGGATCACCAGCATGCTGCAGGCTACCTTCGGCGA
CCTGCTGGCGCCGCGCATCGTCCCGGAAAGCAATGTGCGCCTGACCGCCAGGGAAACCGA
GATGCTCAAGTGGACCGCGGTGGGCAAGACCTACGGCGAGATCGGCCTGATCCTGTCGAT
CGACCAGCGCACGGTGAAATTCCATATCGTCAATGCGATGCGCAAGCTCAACTCCAGCAA
CAAGGCGGAGGCCACCATGAAGGCCTACGCCATCGGCCTGCTCAACTGAATCGACGCCTC
GTCGCCTAGCGAGGCCGCCGC

Fig. 18E

PA14 PhzR SEQ ID NO:165

PhzR peptide sequence

MHDEREGYLEILSRITTEEEFFSLVLEICGNYGFEFFSFGARAPFPLTAPKYHFLSNYPG
EWKSRYISEDYTSIDPIVRHGLLEYTPLIWNGEDFQENRFFWEEALHHGIRHGWSIPVRG
KYGLISMLSLVRSSESIAATEILEKESFLLWITSMLQATFGDLLAPRIVPESNVRLTARE
TEMLKWTAVGKTYGEIGLILSIDQRTVKFHIVNAMRKLNSSNKAEATMKAYAIGLLNZ

Fig. 18F

34H4 SEQ ID NO:118

```
  1  ACCAACATCC  TGGTCCTGAG  CAACAGCCAG  CGCCACGGCC  TGGCCGCCGC
 51  CTGGCCGATC  GTGCTCGGCG  CCTGCGCGGC  GGTGGCGGCG  CTGATCCTGC
101  TGCTCGGGCT  CGGCCTGGGC  GAGCTGCTGC  GGCGCCACCC  GTTGCTCCAG
151  CAGGGGCTCG  CCTGGCTTGG  CGTCGGCTGG  CTCAGCTACC  TGGCCTGGAG
201  CCTGTTCCGC  AGCGCG
```

Fig. 19

33C7 SEQ ID NO:119

```
  1  CCACCGAAGT  AACGGGTCAG  CTCGTCGCAC  AACAGGCGTC  GCTCCTCGGC
 51  CTGCATCAGG  CTGCCCAGCG  GGCCCTGGAA  CCAGTCGCGC  GCGCCCGGTT
101  GAT
```

Fig. 20

25a12.3 SEQ ID NO:120

```
  1  GCGGTGCCCT  GGATGTCGTC  GTTGAAGCAG  CACAGCTCGT  CCTTGTAGCG
 51  CTCCAGCAAC  GGCATGGCAT  TGGTCTGGGC  GAAGTCCTCG  AATTGCAGCA
101  GGACCTTGGG  CCACGGCGCT  TGATCGCCTG  GATGAACAGG  TCGACAA
```

Fig. 21

8C12 SEQ ID NO:121

```
  1  TATTTGTGTA  TAAGNCTCAG  GcTCtGGAGG  GGCCGCTGGG  CAGGCNNAAC
 51  NNCCTCGCGT  NCTNGGCGAC  GANTTNCNNA  TGCTTCGCNT  GCTGCCGGCG
101  TCTCNNCCCT  CNGTACTAgT  CTACGCGTGG  ACAACGTGGC
```

Fig. 22

2A8 SEQ ID NO:122

```
  1  NATTTGTGTA  TAAGAGTCAG  GATCGAACGC  TTCTCTTCGC  CGCAGGAAAG
 51  CCACCGCCGA  GCTGCTGAAG  ATGCTCGAGC  GCAAGGGACA  AGATCATGGG
101  CTTCGGGCAT  NCCNTCTNNA  TCGATTCCTN  CCCACGCAAC  GAAgTGATCA
151  AGGGTTGGTC  GAAGCAGCTC  GCCGACgAGG  TCGGCGACAA  GGTCCTGTTC
201  GCGGTTTCCG  AGGCCATCGA  CAAGACCATG  TGGGAGCAGA  AGAACTGTTC
251  CCCAACGCCG  ACTTCTACCA  CGCCTCGGCG  TCNCCCNTCC  NGTGCTTCCA
301  CCTT
```

Fig. 23

41A5 SEQ ID NO:123
```
  1 tcgttgtaca ggccgaacag gccgagctgc caggtgtcgc cctcg
```

Fig. 24A

50E12 SEQ ID NO: 124
```
  1 gagcagacct gggtacccat ggcttccttg acccgctgca cgatgatgcc cagcgccgcc
 61 ttcagatcct tggcggagtt ctcttcctgg acgatcttgc gcagcgtgtt gagcatgctc
121 ggggccttgt ctccgtgttc agtcccgcgc cagaaggcgc ggggccagtt ccttcagggc
181 gcggcggtag acctcgcgct tgaaggtcac cacctgtccc aggggggtacc agtaactcac
241 ccagcgccag ccgtcgaact cgggcttgct ggtgatatcc atgcgcacgc gcgcctcgtc
301 ggacatcagc cgcagcagga accatttctg cttctggccg atgcacagcg gctggctgtg
361 ggtccgcacc aggcgctgcg gcaaacggta gcgcagccag ccgcgg
```

Fig. 24B

35A9 SEQ ID NO: 125
```
  1 cgcgacagta gcatataatc aatcatgagt gattaattaa ttggcgtttc tgtaacatat
 61 ccttatgatc tgcggcgcct ttcccttgtg aggacgttca gtggccagga aaaccaaaga
121 ggaatcccag aaaacccgcg acggcatact cgatgccgcc gagcgggttt tcctggaaaa
181 gggcgtgggc accactgcca
```

Fig. 24C

pho23 SEQ ID NO: 126
```
  1 tcgatcccaa tgactacaag gacgaaatcc gccagatcgc ccgcgacaag gccaacctgg
 61 agctggacct gaagggcgac atcggctgga gcctgttccc ctggctgggc ctggagc
```

Fig. 24D

6G12 SEQ ID NO: 127
```
  1 ggataggtgc ggcggaaaac gtacgggacg aaagagcggt tttcccgaat gacgcatcct
 61 cctgcaagcg caacttgctg gtggtcgata gcaagtaagg cgcgagacat gtcctgaact
121 tcatgggggc ttttcttat agggcggact gtcgattctg ctagctggta atccttcttt
181 tattgtctct gtgtgcgctt tttgtatgga tgtgtcgaat attttgaata tcgccgttca
241 actttatcca gggccgcagt tcagtgattt attttctcga aaagtttgtt ttttccaata
301 ttcatgcttc atagtctggc cggcc
```

Fig. 24E

25F1 SEQ ID NO: 128
```
  1 gcaggaaacc gttctccana tcctgggcga gaatcctcgg cacatgcacg ccggctccgg
 61 cgagcagtcc ggcgaccttg acgaacggtc ggcagtcttc ctggggcggc ggcgcgtcca
121 tcaccaccag gctgcggtcc cctccctgcc agcggaaata acgacggaag ctggcgtcgc
181 tactggccgg gatcagttcg gcgggggggca cttcccccca accttcggca acgaacaact
241 cgggcaaaca agagtccaac cagcaattca gctgctggaa acgggcatca tcagacattt
301 acggggttct ccacggccct agccgttgcg caggtcatgc tttattatcc agcatctttt
```

Fig. 24F

```
1/1                                               31/11
atg cgt aac ctg att ctc acc gcc atg ctg gcc atg agc ctg ttc
 M   R   N   L   I   L   T   A   M   L   A   M   S   L   F

                         61/21                                 91/31
ggc aag atc gaa gtg gtg gag tac aag gaa tat acc gcc cag gcc gac gac tat acc gcc aag gaa tac gtc gag ctg agc ccg gtg
 G   K   I   E   V   V   E   Y   K   E   Y   T   A   Q   A   D   D   Y   T   A   K   E   Y   V   E   L   S   P   V

121/41                                            151/51
tcc cag ccg ggc aag atc gaa gtg gtg gag ctg ttc tgg tat ggc
 S   Q   P   G   K   I   E   V   V   E   L   F   W   Y   G

                         181/61                                211/71
tgc ccg cat tgc tac gcc ttc gag ccg acc atc gtg ccg tgg agc gag aag ctg ccg gca gat gtc cat ttc gtg
 C   P   H   C   Y   A   F   E   P   T   I   V   P   W   S   E   K   L   P   A   D   V   H   F   V

241/81                                            271/91
cgc ctg cct gcc ctg ttc ggc ggt atc ggc aac gtc cat ggg cag
 R   L   P   A   L   F   G   G   I   G   N   V   H   G   Q

                         301/101                               331/111
atg ttc ctg acc ctg gaa agc atg ggt gtc gag cat gac gtC cac aac gcc gtg ttc gag gcg atc cac aag gag
 M   F   L   T   L   E   S   M   G   V   E   H   D   V   H   N   A   V   F   E   A   I   H   K   E

361/121                                           391/131
cac aag aag ctc gcc act ccg gaa gaG gaa atg gcc gat ttc ctc
 H   K   K   L   A   T   P   E   E   E   M   A   D   F   L

                         421/141                               451/151
gcc ggc aag ggc gtg gac aag gaa aaa ttc ctg agc acc tat aat tcc ttt gcc atc aag ggc cag atg gaa aag gcc
 A   G   K   G   V   D   K   E   K   F   L   S   T   Y   N   S   F   A   I   K   G   Q   M   E   K   A

481/161                                           511/171
aag aag ctg gcg atg gcc tac cag gtc acc ccg acc atg gtg gtc
 K   K   L   A   M   A   Y   Q   V   T   P   T   M   V   V

                         541/181                               571/191
aat ggc aaa tac cgc ttc gac atc ggc tcc gcc ggt ggc ccg gag gaa acc ctc aag ctc gcc gac tac
 N   G   K   Y   R   F   D   I   G   S   A   G   G   P   E   E   T   L   K   L   A   D   Y

601/201
ctg atc gag aaa gag cgc ggc gcc aag aag tag
 L   I   E   K   E   R   A   A   K   K   *
```

Fig. 24G

Sequences of PA14 50E12 encoding for YgdPPa and PtsPpa

```
1/1                          31/11                        61/21                        91/31
GAA AAG GGC CAG ACG CAC GGG GTG ACT CCA TCG GTT GGC GGG TGG CGG gAG GGC CGC GAG AGC CTT TTG CGA AGG CTC CCA CCG GGC CTT GGG AAA aCC CCT AGC CTA CCG GCT TTT GCC


121/41                       151/51                       181/61                       211/71
GGC CCT GTA TCC TCC CCG CAC GAG TCG CAA AGC CGC GCG TTG CCG CTA TCA CAA GCT TTA TGG AAC AAT GCG GGC ACA TGC GAT TTC GAG GAT GTC CCA CTG A  TCG ATT CCG ATG GTT
                                                                                                         M   I   D   S   D   G   F


241/81                       271/91                       301/101                      331/111
TTC GCC CGA ATG TCG GCA TCA TTC TCG CCA ACG AGG CGG GGC AGG TGC TGT GGG CGC GGC GTA TCA ATC AGG AAG CCT GGC AGT TCC CGC AGG GAG GCA TCA ATG ATC GCG AAA CGC CGG
 R   P   N   V   G   I   I   L   A   N   E   A   G   Q   V   L   W   A   R   R   I   N   Q   E   A   W   Q   F   P   Q   G   G   I   N   D   R   E   T   P   E


361/121                      391/131                      421/141                      451/151
AAG AGG CGC TGT ATC GCG AaT TGA ACG AAG AAG TCG GGC TGG AGG CCG GGG ACG TGC GCA TCC TGG CCT GCA CCC GCG GCT GGc TGC GCT ACC GTT TGC CGC AGC GCC TGG TGC GGA CCC
 K   R   R   C   I   A   N   *   T   K   K   S   G   W   R   P   G   T   C   A   S   W   P   A   P   A   A   G   C   A   T   V   C   R   S   A   W   C   G   P


481/161                      511/171                      541/181                      571/191
ACA GCC AGC CGC TGT GCA TCG GCC AGA AGC AGA AAT GGT TCC TGC TGC GGC TGA TGT CCG ACG AGG CGC GCG TGC GCA TGG ATA TCA CCA GCA AGC CGG AGT TCG ACG GcT GGC GCT GGG
 T   A   S   R   C   A   S   A   R   S   R   N   G   S   C   C   G   *   C   P   T   R   R   A   C   A   W   I   S   P   A   S   R   S   S   T   A   G   A   G


601/201                      631/211                      661/221                      691/231
TGA GTT ACT GGT ACC CCC TGG GAC AGG TGG TGA CCT TCA AGC GCG AGG TCT ACC GCC GCG CCC TGA AGg AAC TGG CcC CGc GCC TTC TGG CGC GGG ACT GAA CAC GGA GAC AAG GCC CCG
 S   Y   W   Y   P   L   G   Q   V   V   T   F   K   R   E   V   Y   R   R   A   L   K   E   L   A   P   R   L   L   A   R   D   *


721/241                      751/251                      781/261                      811/271
AGC ATG CTC AAC ACG CTG CGC AAG ATC GTC CAG GAA GTG AAC TCC GCC AAG GAT CTG AAG GCG GCG CTG GGC ATC ATC GTG CAG CGG GTC AAG GAA GCC ATG GGT ACC CAG GTC TGC TCG
 M   L   N   T   L   R   K   I   V   Q   E   V   N   S   A   K   D   L   K   A   A   L   G   I   I   V   Q   R   V   K   E   A   M   G   T   Q   V   C   S


841/281                      871/291                      901/301                      931/311
GTG TAC CTG CTC GAC ACC GAG ACC CAG CGT TTC GTC CTG ATG GCC ACC GAA GGC CTC AAC AAG CGT TCC ATC GGC AAG GTC AGC ATG GCc CCC AGC GAA GGC CTG GTC GGC CTG GTC GGC
 V   Y   L   L   D   T   E   T   Q   R   F   V   L   M   A   T   E   G   L   N   K   R   S   I   G   K   V   S   M   A   P   S   E   G   L   V   G   L   V   G


961/321                      991/331                      1021/341                     1051/351
ACC CGC GAG GAG CCG CTC AAC CTG GAG AAC GCC GCC GCC CAC CCG CGC TAC CGC TAT TTC GCC GAG ACC GGC GAG GAG CGC TAC GCG TCG TTC CTC GGC GCG CCG ATC ATC CAC CAT aGG
 T   R   E   E   P   L   N   L   E   N   A   A   A   H   P   R   Y   R   Y   F   A   E   T   G   E   E   R   Y   A   S   F   L   G   A   P   I   I   H   H   R


1081/361                     1111/371                     1141/381                     1171/391
CGG GTG ATG GGG GTG CTG GTG GTG CAG CAG AAG GAG CGC CGC CAG TTC GAC GAA GGC GAG GAg GCC TTC CTC GTC ACC ATG AGC GCC CAG CTC GCC GGG GTC ATC GCG CAT GCC GAG GCG
 R   V   M   G   V   L   V   V   Q   Q   K   E   R   R   Q   F   D   E   G   E   E   A   F   L   V   T   M   S   A   Q   L   A   G   V   I   A   H   A   E   A


1201/401                     1231/411                     1261/421                     1291/431
ACC GGT TCG ATC CGC GGC CTG GGC AAG CTC GGC AAG GGC ATC CAG GAA GCC AAG TTC GTC GGC GTG CCC GGC GCC CCC GGG GTC GGG GTG GGC AAG GCG GTG GTG GTG TTG CCt CCG GCC
 T   G   S   I   R   G   L   G   K   L   G   K   G   I   Q   E   A   K   F   V   G   V   P   G   A   P   G   V   G   V   G   K   A   V   V   V   L   P   P   A


1321/441                     1351/451                     1381/461                     1411/471
GAC CTG GAA GTG GTG CCG GAC AAG CAG GTC GAC GAC ATC GAC GCC GAG ATC GCC CTG TTC AAG CAG GCC CTG GAG GGC GTT CGC GCC GAC ATG CGC GCG CTG TCG AGC AAG CTC GCC AGC
 D   L   E   V   V   P   D   K   Q   V   D   D   I   D   A   E   I   A   L   F   K   Q   A   L   E   G   V   R   A   D   M   R   A   L   S   S   K   L   A   S


1441/481                     1471/491                     1501/501                     1531/511
CAG tTG CGC AAG GAA GAA CGC GCG CTG TTC GAC GTC TAC CTG ATG ATG CTC GAC GAT GCC TCC ATC GGC AAC GAG GTC AAG CGC ATC ATC CGT ACC GGC CAG TGG GCC CAG GGC GCC CTG
 Q   L   R   K   E   E   R   A   L   F   D   V   Y   L   M   M   L   D   D   A   S   I   G   N   E   V   K   R   I   I   R   T   G   Q   W   A   Q   G   A   L


1561/521                     1591/531                     1621/541                     1651/551
CGC CAG GTG GTG ATG GAG CAC GTG CAG CGC TTC GAG CTG ATG GAC GAC GCC TAT CTC CGC GAG CGC GCC TCC GAC GTC AAG GAC ATc GGT CGC CGC CTG CTC GCC TAC CTg CAG GAA GAa
 R   Q   V   V   M   E   H   V   Q   R   F   E   L   M   D   D   A   Y   L   R   E   R   A   S   D   V   K   D   I   G   R   R   L   L   A   Y   L   Q   E   E


1681/561                     1711/571                     1741/581                     1771/591
CGC AAG CAG AAC CTG ACC TAC CCG GAg CAG ACC ATC ATC GTC AGC GAG GAG CTG TCG CCG GCG ATG CTC GGC GAG GTG CCG GAA GGG CGC CTG GTC GGC CTG GTC TCG GTG CTC GGC TCG
 R   K   Q   N   L   T   Y   P   E   Q   T   I   I   V   S   E   E   L   S   P   A   M   L   G   E   V   P   E   G   R   L   V   G   L   V   S   V   L   G   S


1801/601                     1831/611                     1861/621                     1891/631
GGC AAC TCG CAC GTG GCG ATC CTC GCC CGT GCC ATG GGC ATC CCC ACG GTG ATG GGG GCG GTC GAC CTG CCG TAC TCC AAG GTC GAC GGt ATC GAC CTG ATC GTC GAT GGC TAC CAC GGC
 G   N   S   H   V   A   I   L   A   R   A   M   G   I   P   T   V   M   G   A   V   D   L   P   Y   S   K   V   D   G   I   D   L   I   V   D   G   Y   H   G
```

Fig. 24H

```
1921/641                        1951/651                        1981/661                        2011/671
GAG GTC TAC ACC AAC CCC TCC GCC GAG CTG GTG CGC CAG TAC AGC GAC GTG GTC GCC GAG GAG CGC GAG CTG AGC AAG GGC CTG GCG GCC CTG CGC GAG CTG CCC TGC GAG ACC CTC GAC
E   V   Y   T   N   P   S   A   E   L   V   R   Q   Y   S   D   V   V   A   E   E   R   E   L   S   K   G   L   A   A   L   R   E   L   P   C   E   T   L   D

2041/681                        2071/691                        2101/701                        2131/711
GGC CAC CGC ATG CCG CTC TGG GTC AAC ACC GGC CTG CTC GCC GAT GTC GCC CGC GCC CAG GAG CGT GGC GCC GAG GGC GTG GGC CTG TAC CGC ACC GAA GTG CCG TTC ATG ATC AAC GAC
G   H   R   M   P   L   W   V   N   T   G   L   L   A   D   V   A   R   A   Q   E   R   G   A   E   G   V   G   L   Y   R   T   E   V   P   F   M   I   N   D

2161/721                        2191/731                        2221/741                        2251/751
CGC TTC CCC AGC GAG AAG GAA CAG cTG GCG ATC TAC CGC GAG CAG CTC AGT GCC TTC CAC CCG CTG CCG GTG ACC ATG CGC ACC CTG GAT ATC GGC GGC GAC AAG GCG CTG TCC TAC TTC
R   F   P   S   E   K   E   Q   L   A   I   Y   R   E   Q   L   S   A   F   H   P   L   P   V   T   M   R   T   L   D   I   G   G   D   K   A   L   S   Y   F

2281/761                        2311/771                        2341/781                        2371/791
CCG ATC AAG GAA GAC AAC CCG TTC CTC GGc TGG CGC GGC ATC CGC GTC ACC CTC GAC CAC CCG GAG ATC TTC CTG GTC CAG ACC CGC GCC ATG CTC AAG GCC AGC GAA GGA CTG GAC AAC
P   I   K   E   D   N   P   F   L   G   W   R   G   I   R   V   T   L   D   H   P   E   I   F   L   V   Q   T   R   A   M   L   K   A   S   E   G   L   D   N

2401/801                        2431/811                        2461/821                        2491/831
CTG CGC ATC CTG CTG CCG ATG ATC TCC GGC ACC CAC GAG CTG GAA GAG GCC CTG CAC CTG ATC CAC CGC GCC TGG GGC GAG GTG CGC GAC GAG GGC GTG GAC ATC GCC ATG CCG CCt ATC
L   R   I   L   L   P   M   I   S   G   T   H   E   L   E   E   A   L   H   L   I   H   R   A   W   G   E   V   R   D   E   G   V   D   I   A   M   P   P   I

2521/841                        2551/851                        2581/861                        2611/871
GGC ATG ATG GTC GAG ATT CCC GCC GCC GTG TAC CAG ACC CGC GAG CTG GCC CGt CAG GTC GAC TTC CTT TCG GTC GGT TCG AAC GAC CTG ACC CAG TAC CTG CTG GCG GTC GAC CGC AAC
G   M   M   V   E   I   P   A   A   V   Y   Q   T   R   E   L   A   R   Q   V   D   F   L   S   V   G   S   N   D   L   T   Q   Y   L   L   A   V   D   R   N

2641/881                        2671/891                        2701/901                        2731/911
AAT CCG CGG GTC GCC GAC CTC TAC GAC TAC CTG CAT CCG GCC GTg CTG CAT GCG TTG AAG AAG GTG GTC GAC GAT GCC CAC CTG GAA GGC AAG CCG GTG AGC ATC TGC GGC GAG ATG GCC
N   P   R   V   A   D   L   Y   D   Y   L   E   P   A   V   L   H   A   L   K   K   V   V   D   D   A   H   L   E   G   K   P   V   S   I   C   G   E   M   A

2761/921                        2791/931                        2821/941                        2851/951
GGC GAT CCC GCG GCT GCC GTG CTG CTG ATG GCG ATG GGC TTC GAC AGC CTG TCG ATG AAC GCC ACC AAC CTG CCC AAG GTG AAG TGG CTG CTG CGC CAG ATC ACC CTG GAC AAG GCC CGG
G   D   P   A   A   A   V   L   L   M   A   M   G   F   D   S   L   S   M   N   A   T   N   L   P   K   V   K   W   L   L   R   Q   I   T   L   D   K   A   R

2881/961                        2911/971                        2941/981                        2971/991
GAC CTG CTC GGC CAG TTG CTC ACC TTC GAC AAC CCG CAG GTC ATC CAC AGC TCG CTG CAC CTG GCG TTG CGC AAC CTC GGC CTG GGT CGC GTG ATC AAC CCG GCG GCT ACC GTC CAG CCC
D   L   L   G   Q   L   L   T   F   D   N   P   Q   V   I   H   S   S   L   H   L   A   L   R   N   L   G   L   G   R   V   I   N   P   A   A   T   V   Q   P

3001/1001
TGA TTT TCC C
*
```

Fig. 24I

Sequence of PA14 35A9 encoding mtrRPα

1/1   31/11   61/21   91/31
GTC GAT TTG GAA CAG CAC GGT GCC GGC GCG GAC TgC CTG GCC TTC CTC GTA CAG GCG ACG GGT GAC GAT GCC GGC GAC GCG CGC CCG CgC cTC gGC CTG GCG GTA CGC TTC CAG GCG TCC

121/41   151/51   181/61   211/71
GGG CAG CTC GCT GGT GAT GCC GAT gGG CGC CGG CCT GGC GAC GAT CAC GCC GAC CTC GGC GGG GGC CTC CGC AGT CTT CCC GGt GTC CGC TGC TTC TTC GCA GCC CAG CAG GAA TAG GGC

241/81   271/91   301/101   331/111
GAC CAG GGC CGC CAG CAG CCC GCG CAG CGA GCC GGT CCA TTG GAT GTG CAT GGG TGT CCC TCG ATT CGT GAA CTC GCG AGC TTG CCC GGG AAg GGG CAC CGC AAC TCA CGA GCG GCG CGA

361/121   391/131   421/141   451/151
CAG TAG CAT ATA ATC AAT CAT GAG TGA cTA ATT AAT TGG CGT TTC TGT AAC ATA TCC TTA TGA TCT GCG GCG CCT TTC CCT TGT GAG GAC GTT CAG TGG CCA GGA AAA CCA AAG AGG AAT
                                                                                                                        M   A   R   K   T   K   E   E   S

481/161   511/171   541/181   571/191
CCC AGA AAA CCC GCG AtG GCA TAC TCG ATG CCG CCG AGC GGG TTT TCC TGG AAA AGG GCG TGG GCA CCA CTG CCA TGG CCG ACC TGG CGG ACG CCG CCG GGG TTT CTC GCG GTG CGG TCT
Q   K   T   R   D   G   I   L   D   A   A   E   R   V   F   L   E   K   G   V   G   T   T   A   M   A   D   L   A   D   A   A   G   V   S   R   G   A   V   Y

601/201   631/211   661/221   691/231
ACG GCC ACT ACA AGA ACA AGA TCG AGG TCT GtC TGG CGA TGT GCG ACC GCG CCT TCG GCC AGA TCG AGG TAC CCG AtG AAA ACG CCA GGG TGC CGG CGC TGG AcA TCC TCC TGC GCG CCG
G   H   Y   K   N   K   I   E   V   C   L   A   M   C   D   R   A   F   G   Q   I   E   V   P   D   E   N   A   R   V   P   A   L   D   I   L   L   R   A   G

721/241   751/251   781/261   811/271
GCA TGG GCT TTC TCC GCC AGT GCT GCG AaC CCG GTT CGG TGC AGC GGG TGC TGG AGA TCC TCT ACC TCA AGT GCG AAC GCA GCG ACG AGA ACG AGC CGC TGT TGC GCC GCC GCG AGC TGC
M   G   F   L   R   Q   C   C   E   P   G   S   V   Q   R   V   L   E   I   L   Y   L   K   C   E   R   S   D   E   N   E   P   L   L   R   R   R   E   L   L

841/281   871/291   901/301   931/311
TCG AGA AGC AGG GGC AAC GCT TCG GCC gaC GGC AGA TCC GCC GGG CGG TGG AgC GCG GCG AAC TGC CGG CGC GGC TGG ACG TCG AGC TGG CCA GCA TCT ATC TGC AAT CGC TgT GGG ACG
E   K   Q   G   Q   R   F   G   R   R   Q   I   R   R   A   V   E   R   G   E   L   P   A   R   L   D   V   E   L   A   S   I   Y   L   Q   S   L   W   D   G

961/321   991/331   1021/341   1051/351
GCA TCT GCG GCA CCC TGG CCT GGA CCG AGC GCt TGC GCG ACG ATC CCT GGA gCC GCG CCG AAC GCA TGT TCC GCG CCG GCC TCG AtA GCC TGC GCA GTT CTC CCT ACC TcT TGC TGG CGG
I   C   G   T   L   A   W   T   E   R   L   R   D   D   P   W   S   R   A   E   R   M   F   R   A   G   L   D   S   L   R   S   S   P   Y   L   L   L   A   D

1081/361   1111/371   1141/381   1171/391
ACG CCT GAG GGC GTC AAT CGT CCG CCA TCA GGT GCC TGC GCT GGT CCT CGG CGC CGG CGA CCA CCA GCC GCT GGG CGT CcT CCT CGC TGA TGT GCA GGC GCT TGC CaT CGA TGT AGA GCA
A   *

1201/401   1231/411   1261/421   1291/431
CCG ACA GGC GCG CCT CGG CGT CGG TAC CGA TGC GCA GGC TGT CGA CCG GCG CGC GAT GCC GGC TGC CTT CGA TCT CCA CGC TGC AGA tGC CTT GTT CCG AAT CGA TTT CGA TGG ACA TGG

1321/441   1351/451   1381/461   1411/471
GAa CCT CCc GTT TTc TCC GCC TAC CTT GGG TGG ACC CCG GGC ATC CGC GCG GGT TCT GTC ACG GTA GCT TCA CGC CAG CGT CAC GCG CCT GCC ACC GCG CTT GGC TGC AAT CGT CCG CAG

1441/481   1471/491   1501/501
AGA aGG CGA GGC CAG CGG AGG ACG ACG CCA TGC GGC TAT GCG TGA TTG GTG CGG GCT ATG TGG GAC TGG TGA

Fig. 24J

Sequences of PA14 25F1 encoding for orfT, OrfU and DjlAPa

```
1/1                              31/11                            61/21                            91/31
CGA GGA ATC CAG TCG AGG TGC GAg TAG TCC GCA CTG CGG GAT cTC AGC GCG CGA CCa CCG GAC TCG GTG ACC AGG CGC TGG GTC GGC TCT GCC TCG ACG GTT TCG CCT CCG CTG CCG GAC

121/41                           151/51                           181/61                           211/71
ACG CTG CTG CCC GCC GCG GCG GTG CTG ACC GAG GTC GCG GTA TGC GCC GGG CGC GGT GGC AGG TTG GCA TTG GCG TTC TGC AGC GGG GAG CAA TCC CAG CCG CCG GTG GCC GAT ACC TTG

241/81                           271/91                           301/101                          331/111
CAG TCG AAC TGA TCG GCG GCC TGT ACA GTC AAT GCT GCG ACC GGC TGC AGA GCC AGC AGG CTG CCG GTG ACC AGC AGG GGA AAC TTT CTT CGA AAC ACG AGG GAT TTC ACT GCC ATC TTG

361/121                          391/131                          421/141                          451/151
TTA ATC CGG GCT TCC TGC GCG CCA TCG GCC CGG TGG GCC GCA CGC CTC TCG ATG GGC TGA AAA AGA TGC TGG ATA ATA AAG CAT GAC CTG CGC AAC GGC TAG GGC CGT GGA GAA CCC CGT

481/161                          511/171                          541/181                          571/191
AAA TGT CTG ATG ATG CCC GTT TCC AGC AGC TGA ATc GCT GGT TGG ACT CTT GTT TGC CCG AGT TGT TCG TTG CCG AAG GTT GGG GGG AAG TGC CCC CCG CCG AAC TGA TCC CGG CCA GTA
    M   S   D   D   A   R   F   Q   Q   L   N   R   W   L   D   S   C   L   P   E   L   F   V   A   E   G   W   G   E   V   P   P   A   E   L   I   P   A   S   S

601/201                          631/211                          661/221                          691/231
GCG ACG CCA GCT TCC GTC GTT ATT TCC GCT GGC AGG GAG GGG ACC GCA GCC TGG TGG TGA TGG ACG CGC CGC CGC CCC AGG AAG ACT GCC GAC CGT TCG TCA AGG TCG CCG GAC TGC TCG
    D   A   S   F   R   R   Y   F   R   W   Q   G   G   D   R   S   L   V   V   M   D   A   P   P   P   Q   E   D   C   R   P   F   V   K   V   A   G   L   L   A

721/241                          751/251                          781/261                          811/271
CCG GAG CCG GCG TGC ATG TGC CGA GGA TTC TCG CCC AGG AtC TGG AGA ACG GTT TCC TGC TGC TCA GTG ACC TGG GCC GGC AGA CCT ACC TCG ACG TGC TTC ATC CCG GaA ATG CCG ACG
    G   A   G   V   H   V   P   R   I   L   A   Q   D   L   E   N   G   F   L   L   L   S   D   L   G   R   Q   T   Y   L   D   V   L   H   P   G   N   A   D   E

841/281                          871/291                          901/301                          931/311
AGC TGT TCG AAC CGG CCC TGG ATG CGC TGA TCG CCT TCC AGA AGG TCG ATG TCG CCG GTG TCC TGC CTG CCT ACG ACG AAG CGG TGC TGC GCC GCG AGC TGC AGC TGT TCC CCG ACT GGT
    L   F   E   P   A   L   D   A   L   I   A   F   Q   K   V   D   V   A   G   V   L   P   A   Y   D   E   A   V   L   R   R   E   L   Q   L   F   P   D   W   Y

961/321                          991/331                          1021/341                         1051/351
ACC TGG CCC GCC ACC TCG GCG TGG AGC TGG AGG GCG AGA CGC TGG CCC GCT GGc AgC GGA TCT GCG ACC TGC TGG TAC GCA GCG CGC TGG AGC AAC CGC GGG TGT TCG TCC ATC GCG ACT
    L   A   R   H   L   G   V   E   L   E   G   E   T   L   A   R   W   Q   R   I   C   D   L   L   V   R   S   A   L   E   Q   P   R   V   F   V   H   R   D   Y

1081/361                         1111/371                         1141/381                         1171/391
ATA TGC CGC GCA AcC TGA TGC TCA GCG AGC CCA ACC CGG GCG TCC TCG ACT TCC AGG ACG CCC TGC ACG GCC CGG TCA CCT ACG ATG TCA CCT GCC TGT ACA AGG AtG CCT TCG TCA GTT
    M   P   R   N   L   M   L   S   E   P   N   P   G   V   L   D   F   Q   D   A   L   H   G   P   V   T   Y   D   V   T   C   L   Y   K   D   A   F   V   S   W

1201/401                         1231/411                         1261/421                         1291/431
GGC CGG AGC CGC GCG TGC ATG CCG CGC TGA gtC GTT ACT GGA AGA AGG CGA CCT GGG CCG GCA TCC CGC TGC CGC CAA GCT TCG AgG ACT TCC TCC GcG CCA GCG ACC TGA TGG GCG TGC
    P   E   P   R   V   H   A   A   L   S   R   Y   W   K   K   A   T   W   A   G   I   P   L   P   P   S   F   E   D   F   L   R   A   S   D   L   M   G   V   Q

1321/441                         1351/451                         1381/461                         1411/471
AGC GCC ACC TGA AGG TGA TTG GCA TCT TCG CCC GTA TtT GTC ACC GCG ACG GCA AGC CGC GCT ACC TGG GTG ACG TGC CaC GCT TCT TCC GTT ATC TGG AAA CCG CCG TGG CGC GCC GTC
    R   H   L   K   V   I   G   I   F   A   R   I   C   H   R   D   G   K   P   R   Y   L   G   D   V   P   R   F   F   R   Y   L   E   T   A   V   A   R   R   P

1441/481                         1471/491                         1501/501                         1531/511
CCG AGC TGG CCG AAC TGG GCG AGC TGC TGG CCT CGC TGC CGC AGG GAG CCG AGG CAT GAA GGC GAT GAT CCT CGC CGC CGG cCG TGG CGA GCG CAT GCG GCC GAC CAC CCT GCA CAC GCC
    M   K   A   M   I   L   A   A   G   R   G   E   R   M   R   P   T   T   L   H   T   P   E   L   A   E   L   G   E   L   L   A   S   L   P   Q   G   A   E   A

1561/521                         1591/531                         1621/541                         1651/551
CAA GCC GCT GAT CGA GGC CGC CGG CGT GCC ATT GAT CGA GCG TCA GTT GCT GGC GCT GCG CCA GGC CGG AGT CGA CGA CTG GGT GAT CAA CCA TGC CTG GCT GGG CGA GCA GAT CGA GGC
    K   P   L   I   E   A   A   G   V   P   L   I   E   R   Q   L   L   A   L   R   Q   A   G   V   D   D   W   V   I   N   H   A   W   L   G   E   Q   I   E   A
```

Fig. 24K

1681/561   1711/571   1741/581   1771/591
CTA TCT CGG CGA CGG cTC GCG CCT GGG CGG GCG GAT CGC CTA TTC aCC cGA GGG aGA ACC GCT GGA AAC CGG cGG tGG AAT CTT CCG CGC CCT GCC GTT GCT CGG CGA GCA GCC GTT CCT
 Y   L   G   D   G   S   R   L   G   G   R   I   A   Y   S   P   E   G   E   P   L   E   T   G   G   G   I   F   R   A   L   P   L   L   G   E   Q   P   F   L

1801/601   1831/611   1861/621   1891/631
GTT GCT CAA CGG CGA TGT CTG GAG CGA CTT CGA CTA CTC TCG GCT GCA TCT TGC CGA CGG CGA CCT GGC GCA TCT GGT GCT GGT CGA CAA CCC GGC GCA CCA TCC CGC CGG CGA TTT CCA
 L   L   N   G   D   V   W   S   D   F   D   Y   S   R   L   H   L   A   D   G   D   L   A   H   L   V   L   V   D   N   P   A   H   H   P   A   G   D   F   H

1921/641   1951/651   1981/661   2011/671
CCT GGA TGC CGG CGG ACG GGT GGG CGA GAC CCG CGA AGC GGG CGG CAA CCT gAC CTA CAG CGG GAT CGC CGT ACT GCA TCC CGC GCT GTT CGA GGG CTG CCA GCC GGG CGC CTT CAA GCT
 L   D   A   G   G   R   V   G   E   T   R   E   A   G   G   N   L   T   Y   S   G   I   A   V   L   H   P   A   L   F   E   G   C   Q   P   G   A   F   K   L

2041/681   2071/691   2101/701   2131/711
GGC GCC GCT ATT GCG CAA GGC CAT CGC CGC GGG GCG GGT CAG CGG CGA ACA CtA TCG TGG GCA GTG GGT CGA CGT CGG TAC CCA CGA GCG CCT GGC GGA AGT CGA GCG ATT GCT GGC GGA
 A   P   L   L   R   K   A   I   A   A   G   R   V   S   G   E   H   Y   R   G   Q   W   V   D   V   G   T   H   E   R   L   A   E   V   E   R   L   L   A   E

2161/721   2191/731   2221/741   2251/751
GCA CGC CTG AGA TGC TCT GGC CCG CTA CGC TGA TCG GAG CCG GAG CCG GCT GGG CCC TGG CCA GCA TCC CCG GCG CCC TGC TCG GCG GCC TGC TGG GGC AAC TGC TGG ACC GCA GGT TGC
 H   A   *   M   L   W   P   A   T   L   I   G   A   G   A   G   W   A   L   A   S   I   P   G   A   L   L   G   G   L   L   G   Q   L   L   D   R   R   L   R

2281/761   2311/771   2341/781   2371/791
GCC TGG AGT CCT GGC GCG GCC TGC TGG CGC GCT TGC GCG GGC GGG CGG TGA ACG ATG AGG ACG ACC TGC TGT TcC AGT TGC TCG GCT ATC TGG CCA AGA GCG GCG GGC GGG TGG AGG AGA
 L   E   S   W   R   G   L   L   A   R   L   R   G   R   A   V   N   D   E   D   D   L   L   F   Q   L   L   G   Y   L   A   K   S   G   G   R   V   E   E   N

2401/801   2431/811   2461/821   2491/831
TGC ATA TCC GCC AGG CGC GCG AGG AGA TGG CGT TGC GCA AGC TCG ATA GGC GAG CCC AGC GGC GTG CCA TCG CGT CCT TCG GCA AGG GCA AGG CCG GCA TCG CCC ATC TGC AGG CGG AGG
 H   I   R   Q   A   R   E   E   M   A   L   R   K   L   D   R   R   A   Q   R   R   A   I   A   S   F   G   K   G   K   A   G   I   A   H   L   Q   A   E   V

2521/841   2551/851   2581/861   2611/871
TCG CGC GTC TGA AGG GCG AAC GTG CGG AGG CAG TAT TGC TCG CCT GCT GGC GGA TGG CCT GGG CTG GCG GCG TGC TCA GCC AGT CGG CGC GAC AAC TGG TGT TGC AAT GGG GGC GCT GGC
 A   R   L   K   G   E   R   A   E   A   V   L   L   A   C   W   R   M   A   W   A   G   G   V   L   S   Q   S   A   R   Q   L   V   L   Q   W   G   R   W   L

2641/881   2671/891   2701/901   2731/911
TGG GTT GGT CGG CGG AGC GAA CGG AAC GCT TGT CGG CGC GGG TCA TGC CGA AGC GGA CGC GCG CTG TCG CCC GGG ATA GCT ACC GTG AGG CCC TGC TGC TGC TCG GCG TGG AGG CCG GAA
 G   W   S   A   E   R   T   E   R   L   S   A   R   V   M   P   K   R   T   R   A   V   A   R   D   S   Y   R   E   A   L   L   L   L   G   V   E   A   G   S

2761/921   2791/931   2821/941   2851/951
GCG AGC CGG CGC TGA TCA AAC GCG CCT ATC GCA AGC TGA TCA GCC AGC ATC ATC CGG ACA AAC TGG CGG GAG CCG GCG CCA GCG TCG AGC GCG TGC GTG CGG CTA CCG AGA AAA CCC GTG
 E   P   A   L   I   K   R   A   Y   R   K   L   I   S   Q   H   H   P   D   K   L   A   G   A   G   A   S   V   E   R   V   R   A   A   T   E   K   T   R   E

2881/961   2911/971   2941/981   2971/991
AAT TGC AGG CGG CCT ACG CCC TGG TCC GAG AGC GTG AGG GGT TCC GCT GAT CAC TCC GCA GGT TTC TGC GCA TCG GCC TGC AGG TGA AGA CTG AGC CAG CCG CGG ATT CGT CGG TAC AGT
 L   Q   A   A   Y   A   L   V   R   E   R   E   G   F   R   *

3001/1001   3031/1011   3061/1021   3091/1031
TGC TCC TGC TCC GCC TTG GGG TCG GCC GGT AGa GCC TGC ATC GCG ATT TGT ACG TAG GCc GGG TGT TTC TGC CGC TTG CCG GCC TGC ATG CGC AGC CTG GCC GCC TCG CGG TCG GCG CG

Fig. 24L

PhnA and PhnB SEQ ID NO: 129

```
   1 CTGCAGCGTC TGCCGACCCT GCTGCAACTG ATCCCGGGAC ACGGCGGCCT
  51 GCTGCGGGGG CGGCTGGCCG CGGATGGGGC CGAGTCGGCC TATACCGAGT
 101 GTCTGCGCCT GTGCCGACGG TTGCTCTGGC GCCAGTCCAT GGGCGAGTCC
 151 CTCGACGAAC TGAGCGAGGA GCTGCACCGC GCCTGGGGAG GGCAGAGCGT
 201 CGACTTCCTG CCCGGCGAAC TGCACCTGGG GAGCATGCGC CGGATGCTGG
 251 AGATTCTCTC CCGCCAGGCG CTGCCTCTGG ACTGAGGCGG AACATCCATT
 301 GCGGCGATCG CGCCCGACGG CTGCGGTCGC AATTGGGGGA AATGGGGGTA
 351 TCGATGATGA ATATGCCGTT GCGCGCTAGC GTCGCGCAGG CCAGTCGCCC
 401 ATGGGCGCGG GGAGGTGGCT CGTGAGTGGG GTTGGCTATC GACTGGAAGA
 451 AAGTCTGGAG TACCGCACGC TGGTGCCGGA GGCGCTGTCG ATCTGGCGCA
 501 TGGCTGGCGC CAACCGGATG CTGTTCGACT GCTTCGACGT GGACAGCAAG
 551 GCTGCGCGGC GTAGCGTGGC GATCCTTTCC AGCTGCCTGC GCATCGAGTG
 601 CTGGGGGCGC GATGTGGTGC TGCGGGCGTT GAACTCCAAC GGACGCGCCT
 651 TGCTGGCGCC GTTGAGCGAG GCCTGTCCGG CCCAGGTCAC CTGCTTGCGT
 701 GACGGCGACA CCCTGCACTG GCGCTTCCCC CCGGAAGAGC CGCATGCGGA
 751 CGAGTGGCGA CGCCTGCATG GCCTGTCCAG CCTGGAGGCG CTGCGCCGCG
 801 TGCTCGGAAC GCTGGGCGAC GCGGAGGGGC CTGCGCTGCT GGGCGGCCTG
 851 TTCAGTTTCG ACCTGGCCGA GCAGTTCGAA CCCTTGCCGG CGCCGGCCGA
 901 ACCTGCGCGG CATTGCCCGG ACTACCTGTT CCTGGTGCCG GAGTTGCTGC
 951 TGGATATCGA TCACCTGGCG CGCCGGACTT CGCTGCAAGC GTTCGTCCAC
1001 GATCCGGCCG GGCACGACCG GTTGGCCGCC AGCCTGCGCC AATGTGCCGA
1051 CGAATTCCAT GGCGCCGTGG AGGAGGCTTC CGAGTCGCCG GTGGCAGGCG
1101 TACGGGCCGG CAACTACCAG GTCGACCTGG ACGATGCGAG CTTTGCCCGC
1151 CAGGTAGAAC GCCTGCAGGC CCACGTGAGG GCCGGCGACG TGTTCCAGAT
1201 CGTACCTTCG CGCAGCTTCA GCATGCCGTG CGCGGACCCC TGGCGGGGCCT
1251 ATCGCCAGTT GTGCCTGCGC AACCCCAGCC CGTACCGCTT CTTCCTCGAT
1301 GCGGGGGACT TCTGCCTGTT CGGCGCTTCG CCGGAGTCGG CATTGAAGTA
1351 CGACGCGGAG AGTCGCGAGG TGGAACTCTA TCCCATTGCC GGCACCCGCC
1401 CGCGCGGATG CGATGCCCGG GGCGCCATCG ATGCGGAACT GGACAATCGC
1451 CTGGAAGCGG AGTTGCGCCT GGATGCCAAG GAGATCGCCG AGCACATGAT
1501 GCTGGTCGAC CTGGCGCGCA CGATCTGGC GCGCGTCTGC CGCAGCGGTA
1551 CCCGGCAGGT GCGCGACATG CTCAAGGTCG ATCGCTACAG CCACGTGATG
1601 CACCTGGTCT CGCGCGTGGC TGGCGAACTG CACGGCGAAC TGGATGCGCT
1651 GCATGCCTAC CGTGCCTGCC TGAACATGGG CACCCTGGTC GGCGCGCCGA
1701 AGGTCCGTGC CATGCAGTTG CTGCGGCAGT ACGAGGATGG CTATCGCGGC
1751 AGCTACGGTG GTGCGATCGG CATTCTCGAC AGCGCCGGCA ACCTCGATAC
1801 CAGCATTGTC ATCCGCTCCG CCGAGGTCCG CGAAGGTATC GCGCGGGTTC
1851 GGGCAGGCGC CGGCGTGGTG CTGGATTCGG ATCCACGGCT GGAGGCCGAG
1901 GAAACCCGCA ACAAGGCGCT GGCGGTGCTG ACCGCCGTGG CCGCTGCCGA
1951 ACGCGAAAGG GGAGAGCGCG ATGCGCATCA CGCTGTTGGA TAACTTCGAT
2001 TCCTTCACCT ACAACCTGGT CGAGCAGTTC TGCCTGCTCG CGCGGAGGT
2051 CCGGGTGATG CGCAACGATA CGCCGTTGCC GACGATCCAG GCGGCATTGC
2101 TGGCCGACGG TTGCGAACTG CTGGTGCTGT CGCCGGGGCC CGGTCGGCCG
```

Fig. 25A

```
2151 GAAGACGCCG GCTGTATGCT GGAATTGCTC GCCTGGGCCC GCGGGCGCTT
2201 GCCGGTGCTC GGCGTCTGCC TCGGCCACCA GGCGCTGGCG CTGGCCGCCG
2251 GTGGCGCGGT GGGCGAGGCG AGGAAGCCGC TGCATGGCAA GAGCACGTCC
2301 CTGCGTTTCG ATCAGCGTCA CCCGCTGTTC GACGGCATCG CTGACCTGCG
2351 CGTCGCGCGC TACCACTCGC TGGTGGTCAG TCGCCTGCCG GAAGGTTTCG
2401 ACTGCCTGGC CGATGCCGAT GGCGAGATCA TGGCGATGGC CGATCCGCGC
2451 AATCGACAGC TGGGCTTGCA ATTCCATCCC GAGTCGATTC TCACCACCCA
2501 CGGCCAGCGT CTGCTGGAGA ACGCTCTACT CTGGTGCGGC GCGTTGGCGG
2551 TCGCGGAGCG CCTTCGGGCC TGAGCGGCGC TGCGCAGTTT CGACCGAGGC
2601 TCGGTTGCCA GGCCGGCGCA TCGTCGAAAC GCTGGCGGCC CAGTTCGCGC
2651 AGGCGCTGGC GGGCGCTTTC GAGAAAGCGA CGGAAGCTGC GCTCGGATTC
2701 CAGCGCGGTG TTGTAGTAGC AATACACCTT GGTGTCGATG CCGCCCGGTT
2751 CGTACAGTTC GCTGAGGACT GCCAGGGTAC CGTTGCGCAG GCGTTCCTcG
2801 ACGAAATAAT GCGGCGaGAT GCCCCATCCG ACGCCGGCTT CCACCAGACG
2851 CAGCATGTCG TCGAAGTTTT CCACGAAGAG CACCTTGTCG CTGACCGGCC
2901 GCAGCAGGTT CGAATGCTGC CCGGAGCGGC TgCCGAGGCT GATCTGCCGG
2951 TAATTGGCCA GGCTCGCGAT GCTGTGCAGG GAGGCATTGC ACAACGGGTG
3001 CTGCGGATGG GCGACGACGA ACGCCTTGGT GTAGCCGAGC ACGCACTGGT
3051 TGAAGCGGGA GATCT
```

Fig. 25B

PhnA protein SEQ ID NO:130

```
  1 MGARRWLVSG VGYRLEESLE YRTLVPEALS IWRMAGANRM LFDCFDVDSK
 51 AARRSVAILS SCLRIECWGR DVVLRALNSN GRALLAPLSE DCPAQVTCLR
101 DGDTLHWRFP QEESHADEWR RLHGLSSLEA LRRVLGTLGD AEGPVLLGGL
151 FSFDLAEQFE PLPAPAEPAR HCPDYLFLVP ELLLDIDHLA RRTSLQAFVH
201 DPAGHDRLAA SLRQCADEFH GAVEEASESP VAGVRAGNYQ VDLDDASFAR
251 QVERLQAHVR AGDVFQIVPS RSFSMPCADP WRAYRQLCLR NPSPYRFFLD
301 AGDFCLFGAS PESALKYDAE SREVELYPIA GTRPRGRDAR GAIDAELDNR
351 LEAELRLDAK EIAEHMMLVD LARNDLARVC RSGTRQVRDM LKVDRYSHVM
401 HLVSRVAGEL HGELDALHAY RACLNMGTLV GAPKVRAMQL LRQYEDGYRG
451 SYGGAIGILD SAGNLDTSIV IRSAEVREGI ARVRAGAGVV LDSDPRLEAE
501 ETRNKALAVL TAVAAAERER GERDAHHAVG
```

Fig. 26

PA14 degP SEQ ID NO:131

```
   1 CGTCCGATTC GGCCTGAGTC TTTCTCTTCC CTCGAACATC ACGGGAGCTG TAGTCGATGC
  61 ATACCCTAAA ACGCTGTATG GCTGCGATGG TGGCCTTGCT GGCCTTGAGC CTGGCGATGA
 121 CGGCCCGGGC AGAACTGCCG GACTTCACGC CTTTGGTCGA CAGGCGTCG CCGGCGGTGG
 181 TGAATATCAG TACGCGGCAG AAGCTGCCGG ATCGCGCCAT GGCGCGCGGG CAGCTGTCGA
 241 TCCCCGACCT CGAAGGGCTG CCGCCGATGT TCCGCGACTT CCTCGAGCGC ACGATCCCGC
 301 AGGTTCCGCG CAATCCGCGC GGCCAGCAGC GCGAGGCGCA ATCGCTGGGC TCCGGCTTCA
 361 TCATCTCCAA CGACGGCTAC ATCCTCACCA ACAATACGT CGTGGCCGAT GCCGACGAGA
 421 TCCTGGTGCG CCTGTCCGAC CGTAGCGAGC ACAAGGCCAA GCTGGTCGGC GCGGACCCGC
 481 GCAGCGACGT GGCGGTGCTG AAGATCGAGG CGAAGAACCT GCCGACCCTG AAACTGGGCG
 541 ATTCGAACAA GCTGAAAGTG GGCGAATGGG TCCTGGCCAT CGGTTCGCCG TTCGGCTTCG
 601 ATCACTCGGT CACCGCCGGT ATCGTCAGTG CCAAGGGGCG TAGCCTGCCG AACGAGAGCT
 661 ACGTACCCTT CATCCAGACC GACGTGGCGA TCAACCCGGG CAACTCCGGC GGTCCGCTGC
 721 TGAACCTGGA GGGCGAAGTG GTCGGCATCA ACTCGCAGAT CTTCACCCGT TCCGGCGGCT
 781 TCATGGGCCT GTCCTTCGCC ATCCCGATCG ATGTCGCGCT GAACGTCGCC GACCAGTTGA
 841 AGAAAGCCGG CAAGGTCAGC CGCGGCTGGC TGGGTGTGGT GATCCAGGAA GTGAACAAGG
 901 ATCTCGCCGA GTCCTTCGGC CTCGACAAGC CGTCCGGCGC GCTGGTGGCG CAGCTGGTGG
 961 AAGACGGTCC GGCGGCCAAG GCGGCCTGC AGGTGGGCGA TGTGATCCTC AGCCTGAACG
1021 GCCAGTCGAT CAACGAGTCC GCCGACCTGC CGCACCTGGT GGGCAACATG AAGCCGGGCG
1081 ACAAGATCAA CCTGGACGTG ATTCGCAACG CCAGCGCAA GTCCTTGAGC ATGGCGGTAG
1141 GCAACCTTCC GGACGACGAC GAGGAAATCG CCTCGATGGG CGCTCCGGGC GCCGAGCGCA
1201 GCAGCAACCG CCTGGGCGTG ACCGTCGCCG ACCTGACCGC CGAGCAGCGC AAGAGCCTGG
1261 ATATCCAGGG CGGCGTGGTG ATCAAGAAG TCCAGGACGG TCCGGCCGCG GTCATCGGCC
1321 TGCGTCCGGG CGATGTCATC ACCCACCTGG ACAACAAGGC GGTGACCTCG ACCAAGATCT
1381 CGCCGACGT GGCCAAGGCC CTGCCGAAGA ACCGTTCGGT TTCGATGCGG GTACTG
```

Fig. 27

PA14 degP protein SEQ ID NO: 132

```
  1 MHTLKRCMAA MVALLALSLA MTARAELPDF TPLVEQASPA VVNISTRQKL
 51 PDRAMARGQL SIPDLEGLPP MFRDFLERTI PQVPRNPRGQ QREAQSLGSG
101 FIISNDGYIL TNNHVVADAD EILVRLSDRS EHKAKLVGAD PRSDVAVLKI
151 EAKNLPTLKL GDSNKLKVGE WVLAIGSPFG FDHSVTAGIV SAKGRSLPNE
201 SYVPFIQTDV AINPGNSGGP LLNLEGEVVG INSQIFTRSG GFMGLSFAIP
251 IDVALNVADQ LKKAGKVSRG WLGVVIQEVN KDLAESFGLD KPSGALVAQL
301 VEDGPAAKGG LQVGDVILSL NGQSINESAD LPHLVGNMKP GDKINLDVIR
351 NGQRKSLSMA VGNLPDDDEE IASMGAPGAE RSSNRLGVTV ADLTAEQRKS
401 LDIQGGVVIK EVQDGPAAVI GLRPGDVITH LDNKAVTSTK IFADVAKALP
451 KNRSVSMRVL
```

Fig. 28

PA 8830 algD SEQ ID NO:133

```
   1 GCGCGACAAA CAATCGAGGT GAATGCGATG CGAATCAGCA TCTTTGGTTT
  51 GGGCTATGTC GGTGCAGTAT GTGCTGGCTG CCTGTCGGCA CGCGGTCATG
 101 AAGTCATTGG TGTGGATGTC TCCAGCACCA AGATCGACCT GATCAACCAG
 151 GGCAAGTCGC CCATCGTCGA ACCGGGCCTG GAAGCGTTGT TGCAGCAAGG
 201 CCGGCAGACC GGACGGCTGT CGGGCACCAC CGACTTCAAG AAGGCTGTGC
 251 TGGACTCCGA CGTATCGTTC ATCTGCGTCG GCACGCCGAG CAAGAAGAAC
 301 GGCGACCTGG ACCTGGGCTA CATCGAGACC GTCTGCCGCG AGATCGGCTT
 351 CGCCATCCGC GAGAAGTCCG AACGCCACAC CGTGGTGGTG CGCAGCACCG
 401 TACTGCCGGG CACCGTCAAC AACGTGGTGA TCCCGCTGAT CGAGGACTGC
 451 TCGGGCAAGA AGGCCGGGGT CGACTTCGGC GTCGGCACCA ACCCCGAATT
 501 CCTCCGCGAG AGCACCGCGA TCAAGGACTA CGACTTCCCG CCGATGACCG
 551 TGATCGGCGA ACTGGACAAG CAGACCGGCG ACCTTCTCGA GGAAATCTAC
 601 CGCGAGCTGG ACGCGCCGAT CATCCGCAAG ACCGTCGAGG TCGCCGAGAT
 651 GATCAAGTAC ACCTGCAACG TCTGGCACGC CGCCAAGGTC ACCTTCGCCA
 701 ACGAGATCGG CAACATCGCC AAGGCGGTCG GCGTCGACGG CCGCGAGGTG
 751 ATGGACGTGA TCTGCCAGGA CCACAAGCTC AACCTGTCGC GCTACTACAT
 801 GCGTCCCGGC TTCGCCTTCG GCGGCTCCTG CCTGCCCAAG GATGTACGCG
 851 CCCTCACCTA TCGCGCCAGC CAGCTGGACG TCGAGCACCC GATGCTCGGT
 901 TCGTTGATGC GCAGCAACTC CAACCAGGTG CAGAAGGCCT TCGATCTCAT
 951 CACCAGCCAC GACACCCGCA AGGTCGGCCT GCTCGGCCTG TCGTTCAAGG
1001 CCGGCACCGA CGATTTGCGC GAAAGCCCGC TGGTGGAGCT GGCCGAGATG
1051 CTCATCGGCA AGGGCTACGA GTTCCGCATC TTCGACCGCA ACGTCGAATA
1101 CGCGCGTGTC CACGGGGCCA ACAAGGAATA CATCGAGTCG AAGATCCCGC
1151 ACGTCTCCTC GCTGCTGGTC TCCGACCTCG ACGAAGTGGT GGCGAGTTCC
1201 GATGTGCTGG TGCTGGGCAA TGGCGACGAG CTGTTCGTCG ACCTGGTGAA
1251 CAAGACCCCG AGCGGCAAGA AGCTGGTCGA CCTGGTGGGC TTCATGCCGC
1301 ACACCACCAC TGCCCAGGCC GAGGGCATCT GCTGGTAGCG G
```

Fig. 29

PA 8830 algD protein SEQ ID NO: 134

```
  1 MRISIFGLGY VGAVCAGCLS ARGHEVIGVD VSSTKIDLIN QGKSPIVEPG
 51 LEALLQQGRQ TGRLSGTTDF KKAVLDSDVS FICVGTPSKK NGDLDLGYIE
101 TVCREIGFAI REKSERHTVV VRSTVLPGTV NNVVIPLIED CSGKKAGVDF
151 GVGTNPEFLR ESTAIKDYDF PPMTVIGELD KQTGDLLEEI YRELDAPIIR
201 KTVEVAEMIK YTCNVWHAAK VTFANEIGNI AKAVGVDGRE VMDVICQDHK
251 LNLSRYYMRP GFAFGGSCLP KDVRALTYRA SQLDVEHPML GSLMRSNSNQ
301 VQKAFDLITS HDTRKVGLLG LSFKAGTDDL RESPLVELAE MLIGKGYEFR
351 IFDRNVEYAR VHGANKEYIE SKIPHVSSLL VSDLDEVVAS SDVLVLGNGD
401 ELFVDLVNKT PSGKKLVDLV GFMPHTTTAQ AEGICW
```

Fig. 30

>Contig1126 of Mutant 25A12 SEQ ID NO: 135

..AACACCGGACGCGCCCCGATCATGTGCGCTGAGCGCTACGCTACCGTCAA
CGAAAAAGGCCACCTCGGGGTGGCCTTTTCGCGTTCTCGCACCGATCGCG
CGGAATATCGGCGGTTAACGCCTCTCCCCCGTGCGCACCTGCGGCTGAGC
CTCAGAACGAAGTCCGGCGGTAGGCACGGTAGCGCGGGAACCAGAAGTTC
GCCTCGATGGCGTCGTTCAGTACCTCGTCGCTGGTATGCAGGGCCTTGCC
CTCGGCCTGGGCCTGCTTGGCCACGGCGACGGCGATGCGCTTGCTGACCT
CGCGGATGTCGCCCAGCGCCGGCAACACGGCGCCCTCGCCCTGGGTAACG
ATCGGCGAGCAGTTGGCCAGGGCGTTGGCCGCGGCCATCAGCATGCCTTC
GGTGACCCGATTGGCCCGCGCGGCGATCACCCCCAGGCCGATGCCGGGGA
AGATATAGGCGTTGTTGCACTGGGCGATGGGAATCCGCTTGTCGCCCACC
TGCACCGGTTGGAACGGGCTACCGGTGGCGACCAGCGCCTGGCCGTCGGT
CCAGTTGAGGATTTCCTGCGGAGTCGCCTCGACCCGCGAGGTCGGGTTGG
ACAGCGGCATCACCAGCGGCTGCTTGCAATGGCTGTGCAGCTCACGGATG
ACCTCTTCGGAAAACAGCCCGCGCTGCCCGGAGACGCCGATCAGCACCGT
CGGCCGGGCATTGCGGATCACTTCCAGCAACGCCAGGTCGTCGCCCTGCT
GGCCGCCCCAGGCACCGAGATCGGCGCGCTTCTGCGCCAGGCGGTGCTGG
AAGTCGACCAGGTTGCTCATGTCGTCGGTGAGCAGGCCCCAGCGGTCGAC
CATGAAGATGCGCCGACGCGCCTGGGCCTCGTCCAGGCCCTCCAGTTGCA
TGGCGGCGATGATCTGTTCGGCGATGCCGCAACCGGCGGAGGGGCGCCGA
CGAAGGTCACGGTCTGCTCGCTGAGCTTCTCGCCCTTGGCCTTGCAAGCC
GCCAGCAGGGTGCCCACGGCCACCGCGGCGGTGCCCTGGATGTCGTCGTT
GAAGCAGCACAGCTCGTCCTTGTAGCGCTCCAGCAACGGCATGGCATTGG
TCTGGGCGAAGTCCTCGAATTGCAGCAGGACGTTGGGCCAGCGGCGCTTG
ATCGCCTGGATGAACAGGTCGACGAACTCCTCGTACTGCGCCCCGCTCAC
CCGCTCGTGGCGCCACCCAATGTACATCGGGTCGTTGAGCAGGTCCGGGT
TGTTGGTGCCGACGTCCAGCACCACCGGCAGGGTGTAGGCCGGGCTGATA
CCGCCGCAGGTGTAACAGGGGACAGCTTGCCGATCGGGATGCCCATCCGG
CCGATGCCCTGGTTGCCGAGGGCGAGGATCGGCTGGCTGTCGGTACAAAA
CAATCTAAGGTGTCTTTGGTGGCTTGAAGGAGTTTCAATCGTTCGGGCCG
GGAAGAATAAAGGCCCGGTGGGTCGAAACTTTGAATCTGGAAGGTTGCAA
ACTGGGGGAAAAAATGGAAATTTTTAAGAGCCTAAGAGCGGAAAAAAGTT
CTTTTTCTAAAAAGAAAAAAATGGGGAAAAAGTTGAAAAGTATATGATAA
GAGCAGGTGTCAAAATGAATGTTTTGAAAGCCCAGTGAAATAAACTCTGG
AAAAGGCAGTTATAAGGGCTATAAAAGGGATGAAAAAAGAAGTGTGTGAA
ATAACGAAAGGCAATAGGGAAAA

Fig. 31

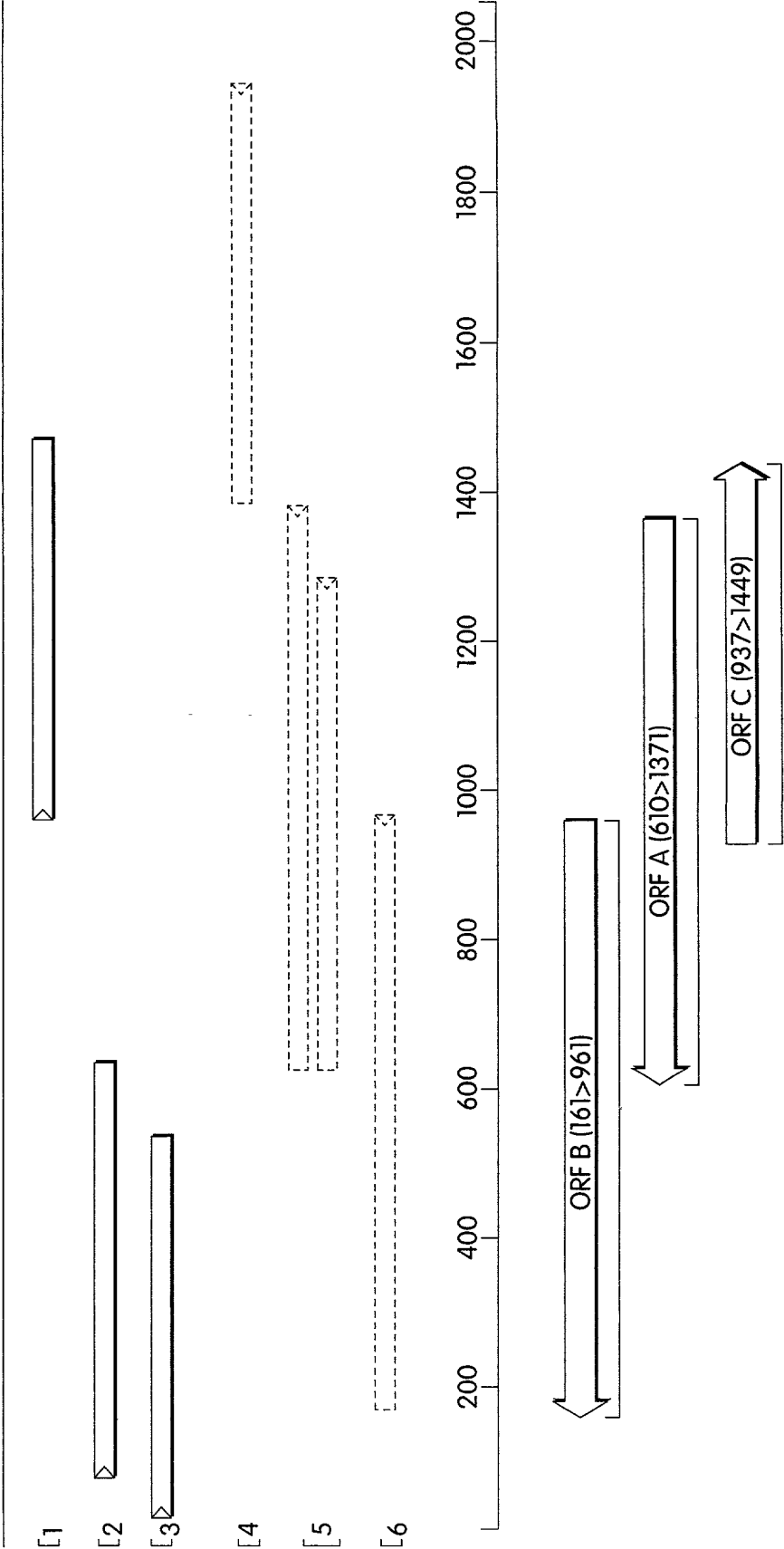33C7 contig Map (1 > 2048) ORF Map

[1

[2

[3

[4

[5

[6

200 400 600 800 1000 1200 1400 1600 1800 2000

ORF B (161>961)

ORF A (610>1371)

ORF C (937>1449)

Fig. 32A

Sequence: 33C7 contig From: 1 To: 2048

```
          10        20        30        40        50        60
AGCTTATGCA TGCGGCCGCA TCTAGAGGGC CCGGATCCGG TGACCATCGG TCACCGGCAT 60
GCCGGTGGTT TCGGTATCCA GTACGACGCT ACGCATCTAT AGAGCCTTTC TCTGTTTCGC 120
TGCAGCCGTG GCTGCTGAAC GCTTGTTTCG GTGTGGCCGC TCAGCGCGGC AATTCGGCGA 180
CGCCACGGTT GGCCAACTGG TCGGCCCGCT CGTTGCCGGG GTCGCCGGTA TGCCCGCGGA 240
CCCACTGCCA CTCCACCTGG TGCCGGGCGA CCTGTTCATC CAGGGCCTGC CAGAGGTCGG 300

         310       320       330       340       350       360
CATTCTTGAC AGGCTGCTTG CTGGCGGTCT TCCAGCCGCG CTTCTTCCAG TTCGGCAACC 360
ATTCGGTGAT GCCGCGCATC ACGTATTCCG AGTCGGTGAT CAGACGGATC GGACAGGAAC 420
GCTTGAGTGC CGCCAGCGCC TGGATCGCCG CCATCAGCTC CATGCGGTTG TTGGTGGTGT 480
CCGGCTCGCC GCCCCAAAGC TCTCGCTCGG CGCCCTTGTA GAGGAGCAAC GCCCCCCAGC 540
CGCCGCGCCC AGGGTTGCCC TTGCAGGCGC CGTCGGTATA GATCACTACC TGTTCTTTAT 600

         610       620       630       640       650       660
CTGTCATGCC TAAATTTCGG AATCTCGCCG GCTGACTTTC GCCACCGGCA TGGGCACCAG 660
CTGACCGCGC GGTTCGCGCT TGCTCTGGCG CAACGGGCGC AACCCCACGA CCAGCTTGCG 720
TGCCACCAAT AGATAGAAGC CGGCGCCCGA AGACTGCCAG GCGTCGCCCC AGCGCTCCAG 780
GCGAGCCAGG CGCGATTGCC AGGCTGCCGA CGCAAGCGGC GGACGATAGC ACCCGAAGCG 840
CCGTTTCTCC AGCGCGAAGC CCAGCAGGTT GAGCCAATCG CAGGCCCGCG ACGGAGGAAT 900

         910       920       930       940       950       960
GCAGCGGGCC TGGCGCAAGG CATCCCCGGC GAAATAATGA CGGATGCCCC ACAGGCTCCA 960
TGGGTTGATG CCGATCAGCA GCAGGTGGCC GCCCGGACGA ACGGTACGCG CGGCTTCGCG 1020
CAGGAGACGG TGAGGCGACA GGCAGAAATC CAGGCCGTGT TGCAGCAGGA CCACGTCCGC 1080
GGCATGTTCG CTGAGCGGCC AGGCGCCCTC TTCGCAGGCG ATGTCCACGC CCGGCAGCGG 1140
CGGCCCCAGG CGCACGCCGC GCTGAATCTG CCCGGTGCTC GGCGGCAGTT CGGCATGCGG 1200

        1210      1220      1230      1240      1250      1260
CCCGTAGTGC ACCAGGTAGC CACCGAAGTA ACGGGTCAGC TCGTCGCACA ACAGGCGTCG 1260
CTCCTCGGCC AGCATCAGGC TGCCCAGCGG GCCCTGGAAC CAGTCGCGCG CCCGGTTGAT 1320
CGATGCCAGC CACTCGGCAT CGGTCTGGGC GAAGGCTTGC GGTTCGTTCA TGCGTACCTC 1380
CAGCGTCTTC CCCTTCGCGG CGACGGACGC CGGCACGACG GGAAAATAAG CAATACTATG 1440
CGCCAATGAC TTCTGCTTAG CGACATCGAC CCATGATACA GATCGACGCC CTGCCCGCCT 1500
```

Fig. 32B

Sequence: 33C7 contig From: 1 To: 2048 (continued)

```
         1510      1520      1530      1540      1550      1560
   ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴
TCAACGACAA CTACATCTGG CTGTTGCAAG ATGCGACAAG CCGTCGCTGC GCGGTGGTCG 1560
ACCCCGGCGA TGCCAAGCCG GTGGAAGCCT GGCTGGCCGC CCATCCCGAC TGGCGGTTGA 1620
GCGATATCCT GGTGACCCAC CACCATCACG ACCACGTCGG CGGCGTCGCG GCCCTGAAGG 1680
AACTGACCGG CGCGCGGGTT CTCGGCCCGG CCAACGAGAA GATCCCGGCC CGCGACCTGG 1740
CGCTGGAAGA CGGCGAACGG GTCGAGGTGC TCGGCCTGGT CTTCGAGATC TTCCACGTGC 1800
         1810      1820      1830      1840      1850      1860
   ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴ ┴┴┴┴┴┴┴┴┴┴┴
CCGGCCATAC CCTCGGCCAT ATCGCCTACT ACCACCCGGC GGAGACGCCG CTGCTGTTCT 1860
GCGGCGACAC CCTGTTCGCC GCCGGCTGCG GCCGTCTCTT CGAAGGCACC CCGGCGCAGA 1920
TGCACCATTC CCTGGCGCGA CTGGCCGCGC TGCCGGCCAA CACCCGGGTC TACTGCACCC 1980
ACGAGTACAC GCTGAGCAAC CTGCGCTTCG CGCTGGCGGT GGAGCCCGAC AACGCGGCGC 2040
TGCGGGAA 2048
```

Fig. 32C

33C7 ORF A

```
ATGAACGAAC CGCAAGCCTT CGCCCAGACC GATGCCGAGT 40
GGCTGGCATC GATCAACCGG GCGCGCGACT GGTTCCAGGG 80
CCCGCTGGGC AGCCTGATGC TGGCCGAGGA GCGACGCCTG 120
TTGTGCGACG AGCTGACCCG TTACTTCGGT GGCTACCTGG 160
TGCACTACGG GCCGCATGCC GAACTGCCGC CGAGCACCGG 200


GCAGATTCAG CGCGGCGTGC GCCTGGGGCC GCCGCTGCCG 240
GGCGTGGACA TCGCCTGCGA AGAGGGCGCC TGGCCGCTCA 280
GCGAACATGC CGCGGACGTG GTCCTGCTGC AACACGGCCT 320
GGATTTCTGC CTGTCGCCTC ACCGTCTCCT GCGCGAAGCC 360
GCGCGTACCG TTCGTCCGGG CGGCCACCTG CTGCTGATCG 400


GCATCAACCC ATGGAGCCTG TGGGGCATCC GTCATTATTT 440
CGCCGGGGAT GCCTTGCGCC AGGCCCGCTG CATTCCTCCG 480
TCGCGGGCCT GCGATTGGCT CAACCTGCTG GGCTTCGCGC 520
TGGAGAAACG GCGCTTCGGG TGCTATCGTC CGCCGCTTGC 560
GTCGGCAGCC TGGCAATCGC GCCTGGCTCG CCTGGAGCGC 600


TGGGGCGACG CCTGGCAGTC TTCGGGCGCC GGCTTCTATC 640
TATTGGTGGC ACGCAAGCTG GTCGTGGGGT TGCGCCCGTT 680
GCGCCAGAGC AAGCGCGAAC CGCGCGGTCA GCTGGTGCCC 720
ATGCCGGTGG CGAAAGTCAG CCGGCGAGAT TCCGAAATTT 760
AG 762
```

Fig. 32D

Sequence: 33C7 ORF A From: 1 To: 254

```
           10          20          30          40
 ||||||||||| |||||||||| ||||||||||| |||||||||||
MNEPQAFAQT DAEWLASINR ARDWFQGPLG SLMLAEERRL 40
LCDELTRYFG GYLVHYGPHA ELPPSTGQIQ RGVRLGPPLP 80
GVDIACEEGA WPLSEHAADV VLLQHGLDFC LSPHRLLREA 120
ARTVRPGGHL LLIGINPWSL WGIRHYFAGD ALRQARCIPP 160
SRACDWLNLL GFALEKRRFG CYRPPLASAA WQSRLARLER 200

WGDAWQSSGA GFYLLVARKL VVGLRPLRQS KREPRGQLVP 240
MPVAKVSRRD SEI. 254
```

Fig. 32E

Sequence: 33C7 ORF B From: 1 To: 801

```
           10          20          30          40          50          60
 ||||||||||| |||||||||| ||||||||||| |||||||||| ||||||||||| ||||||||||
ATGGAGCCTG TGGGGCATCC GTCATTATTT CGCCGGGGAT GCCTTGCGCC AGGCCCGCTG 60
CATTCCTCCG TCGCGGGCCT GCGATTGGCT CAACCTGCTG GGCTTCGCGC TGGAGAAACG 120
GCGCTTCGGG TGCTATCGTC CGCCGCTTGC GTCGGCAGCC TGGCAATCGC GCCTGGCTCG 180
CCTGGAGCGC TGGGGCGACG CCTGGCAGTC TTCGGGCGCC GGCTTCTATC TATTGGTGGC 240
ACGCAAGCTG GTCGTGGGGT TGCGCCCGTT GCGCCAGAGC AAGCGCGAAC CGCGCGGTCA 300
           310         320         330         340         350         360
 ||||||||||| |||||||||| ||||||||||| |||||||||| ||||||||||| ||||||||||
GCTGGTGCCC ATGCCGGTGG CGAAAGTCAG CCGGCGAGAT TCCGAAATTT AGGCATGACA 360
GATAAAGAAC AGGTAGTGAT CTATACCGAC GGCGCCTGCA AGGGCAACCC TGGGCGCGGC 420
GGCTGGGGGG CGTTGCTCCT CTACAAGGGC GCCGAGCGAG AGCTTTGGGG CGGCGAGCCG 480
GACACCACCA ACAACCGCAT GGAGCTGATG GCGGCGATCC AGGCGCTGGC GGCACTCAAG 540
CGTTCCTGTC CGATCCGTCT GATCACCGAC TCGGAATACG TGATGCGCGG CATCACCGAA 600
           610         620         630         640         650         660
 ||||||||||| |||||||||| ||||||||||| |||||||||| ||||||||||| ||||||||||
TGGTTGCCGA ACTGGAAGAA GCGCGGCTGG AAGACCGCCA GCAAGCAGCC TGTCAAGAAT 660
GCCGACCTCT GGCAGGCCCT GGATGAACAG GTCGCCCGGC ACCAGGTGGA GTGGCAGTGG 720
GTCCGCGGGC ATACCGGCGA CCCCGGCAAC GAGCGGGCCG ACCAGTTGGC CAACCGTGGC 780
GTCGCCGAAT TGCCGCGCTG A 801
```

Fig. 32F

Sequence: 33C7 ORF B PROTEIN From: 1 To: 267

```
           10          20          30          40          50
 ||||||||||| |||||||||| ||||||||||| |||||||||| ||||||||||
MEPVGHPSLF RRGCLAPGPL HSSVAGLRLA QPAGLRAGET ALRVLSSAAC 50
VGSLAIAPGS PGALGRRLAV FGRRLLSIGG TQAGRGVAPV APEQARTARS 100
AGAHAGGESQ PARFRNLGMT DKEQVVIYTD GACKGNPGRG GWGALLLYKG 150
AERELWGGEP DTTNNRMELM AAIQALAALK RSCPIRLITD SEYVMRGITE 200
WLPNWKKRGW KTASKQPVKN ADLWQALDEQ VARHQVEWQW VRGHTGDPGN 250
           260         270         280         290         300
 ||||||||||| |||||||||| ||||||||||| |||||||||| ||||||||||
ERADQLANRG VAELPR. 267
```

Fig. 32G

33C7 ORF C

```
         10        20        30        40        50        60
     |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
ATGACGGATG CCCCACAGGC TCCATGGGTT GATGCCGATC AGCAGCAGGT GGCCGCCCGG 60
ACGAACGGTA CGCGCGGCTT CGCGCAGGAG ACGGTGAGGC GACAGGCAGA AATCCAGGCC 120
GTGTTGCAGC AGGACCACGT CCGCGGCATG TTCGCTGAGC GGCCAGGCGC CCTCTTCGCA 180
GGCGATGTCC ACGCCCGGCA GCGGCGGCCC CAGGCGCACG CCGCGCTGAA TCTGCCCGGT 240
GCTCGGCGGC AGTTCGGCAT GCGGCCCGTA GTGCACCAGG TAGCCACCGA AGTAACGGGT 300
         310       320       330       340       350       360
     |||||||||| |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
CAGCTCGTCG CACAACAGGC GTCGCTCCTC GGCCAGCATC AGGCTGCCCA GCGGGCCCTG 360
GAACCAGTCG CGCGCCCGGT TGATCGATGC CAGCCACTCG GCATCGGTCT GGGCGAAGGC 420
TTGCGGTTCG TTCATGCGTA CCTCCAGCGT CTTCCCCTTC GCGGCGACGG ACGCCGGCAC 480
GACGGGAAAA TAAGCAATAC TATGCGCCAA TGA 513
```

Fig. 32H

Sequence: 33C7 ORF C PROTEIN From: 1 To: 171

```
         10        20        30        40        50
     |||||||||| |||||||||| |||||||||| |||||||||| ||||||||||
MTDAPQAPWV DADQQQVAAR TNGTRGFAQE TVRRQAEIQA VLQQDHVRGM 50
FAERPGALFA GDVHARQRRP QAHAALNLPG ARRQFGMRPV VHQVATEVTG 100
QLVAQQASLL GQHQAAQRAL EPVARPVDRC QPLGIGLGEG LRFVHAYLQR 150
LPLRGDGRRH DGKISNTMRQ . 171
```

Fig. 32I

1G2 SEQ ID NO:137

```
  1 NTTGTGTTAA  GATCAGGCTT  GGTGGTGAAG  AAAGGTTCGA  ACNNGTGGTC
 51 AATGATCNAC  TTCGGGGATN  CNGCTGCCCG  TATNATTCAA  CACGTGGTCA
101 AACGGTATGT  TCCGAGGCGT  CTGNCCACCN  GTACTAGTCG  ACGC
```
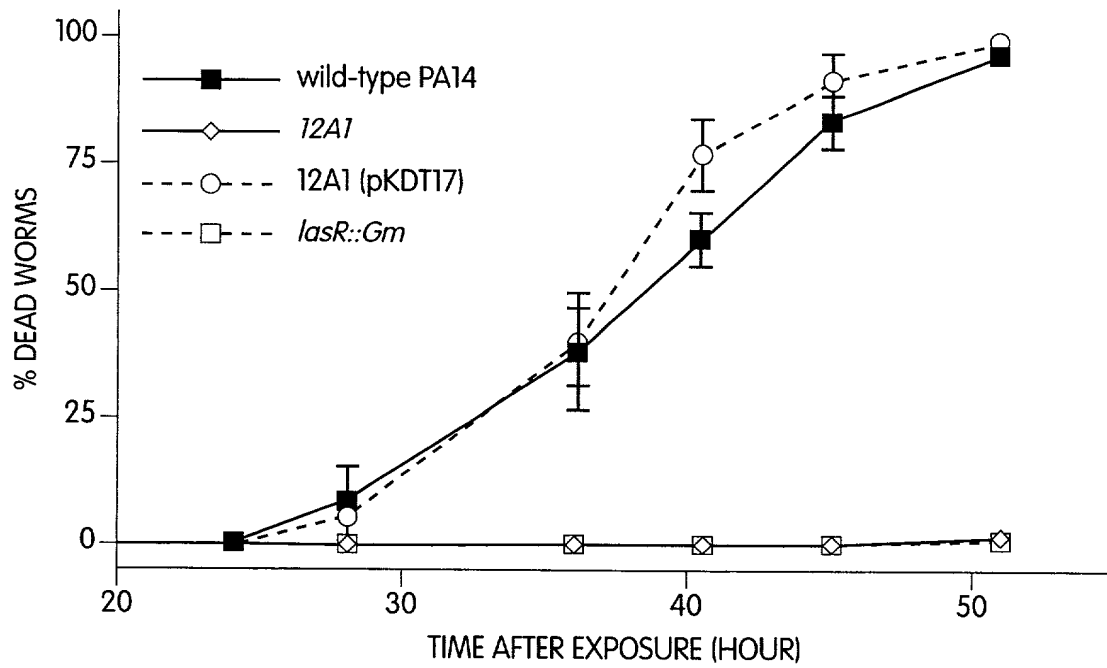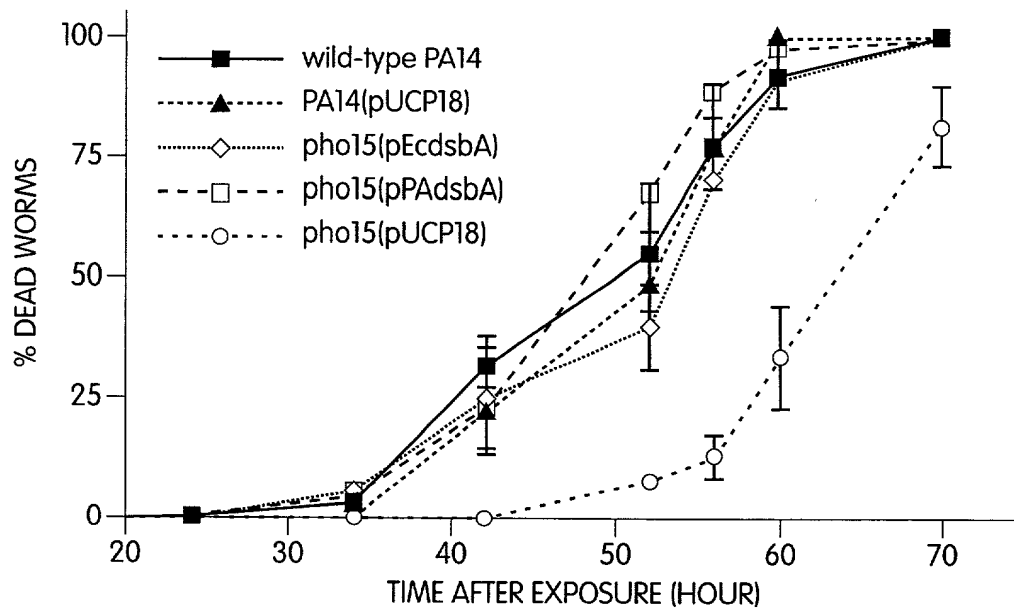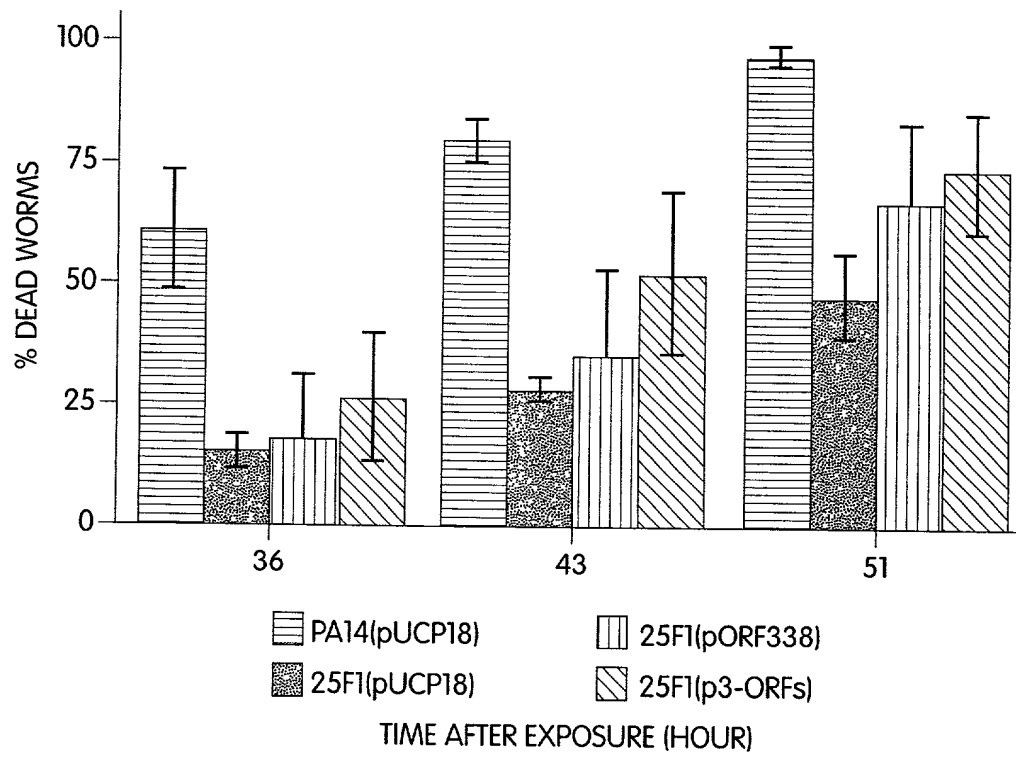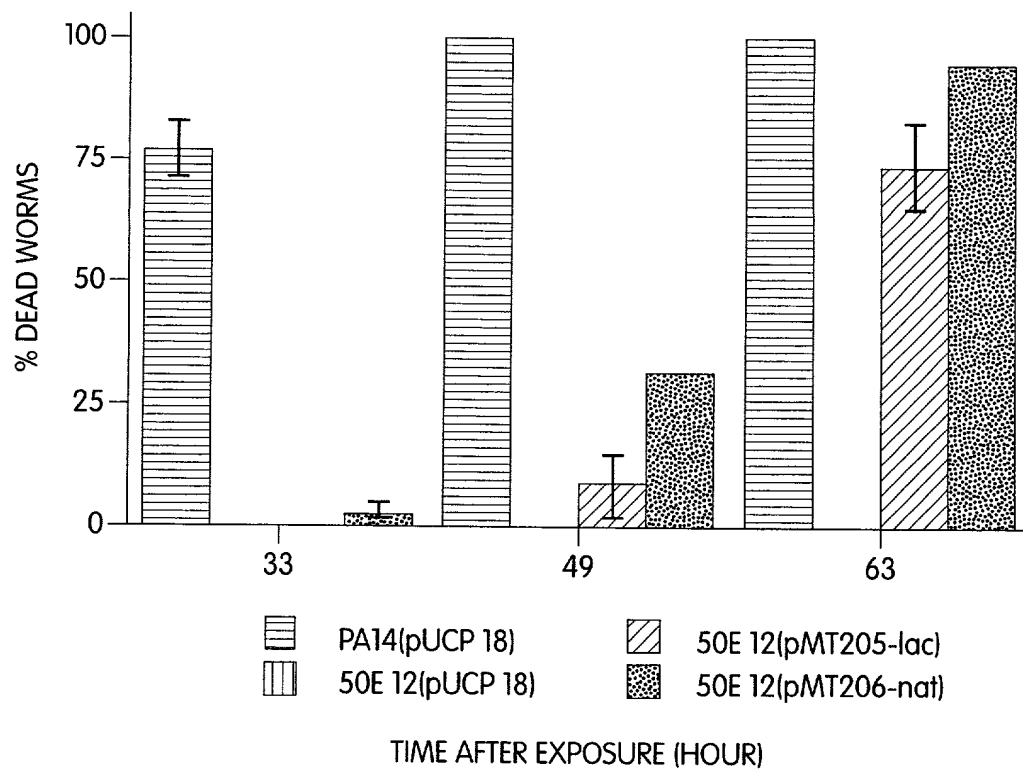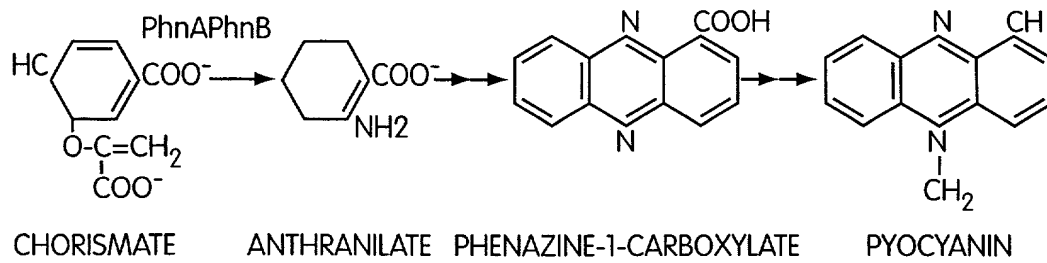
Fig. 33

Fig. 34A

Fig. 34B

Fig. 34C

Fig. 34D

Fig. 35A

Fig. 35B

Fig. 35C